Title

What Makes Thienoguanosine an Outstanding Fluorescent DNA Probe?

Journal

Journal of the American Chemical Society, 142(40)

ISSN

0002-7863

Authors

Kuchlyan, Jagannath
Martinez-Fernandez, Lara
Mori, Mattia
et al.

Publication Date

2020-10-07

DOI

10.1021/jacs.0c06165

Peer reviewed

HHS Public Access

Author manuscript

J Am Chem Soc. Author manuscript; available in PMC 2021 October 07.

Published in final edited form as:

J Am Chem Soc. 2020 October 07; 142(40): 16999–17014. doi:10.1021/jacs.0c06165.

What makes thienoguanosine an outstanding fluorescent DNA probe?

Jagannath Kuchlyan[+,a], **Lara Martinez-Fernandez**[+,b], **Mattia Mori**[+,c], **Krishna Gavvala**[a], **Stefano Ciaco**[a,c], **Christian Boudier**[a], **Ludovic Richert**[a], **Pascal Didier**[a], **Yitzhak Tor**[d], **Roberto Improta**[b], **Yves Mély**[*,a]

[a]Laboratoire de Bioimagerie et Pathologies, UMR 7021 CNRS Université de Strasbourg, Faculté de pharmacie, 74 route du Rhin, 67401 Illkirch, France.

[b]Departamento de Química, Facultad de Ciencias and Institute for Advanced Research in Chemistry (IADCHEM), Universidad Autónoma de Madrid, Campus de Excelencia UAM-CSIC, Cantoblanco, 28049 Madrid, Spain.

[c]Department of Biotechnology, Chemistry and Pharmacy, University of Siena, Via Aldo Moro 2, 53100 Siena, Italy.

[d]Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA, 92093-0358, USA.

Abstract

Thienoguanosine (thG) is an isomorphic guanosine (G) surrogate that almost perfectly mimics G in nucleic acids. To exploit its full potential and lay the foundation for future applications, twenty DNA duplexes, where the bases facing and neighboring thG were systematically varied, were thoroughly studied using fluorescence spectroscopy, molecular dynamics simulations and mixed quantum mechanical/molecular mechanics calculations, yielding a comprehensive understanding of its photophysics in DNA. In matched duplexes, thG's hypochromism was larger for flanking G/C residues but its fluorescence quantum yield (QY) and lifetime values were almost independent of the flanking bases. This was attributed to high duplex stability, which maintains a steady orientation and distance between nucleobases, so that a similar charge transfer (CT) mechanism governs the photophysics of thG independently of its flanking nucleobases. thG can therefore replace any G residue in matched duplexes, while always maintaining similar photophysical features. In contrast, the local destabilization induced by a mismatch or an abasic site restores a strong dependence of thG's QY and lifetime values on its environmental context, depending on the CT route efficiency and solvent exposure of thG. Due to this exquisite sensitivity, thG appears ideal for monitoring local structural changes and single nucleotide polymorphism. Moreover, thG's dominant fluorescence lifetime in DNA is unusually long (9–29 ns), facilitating its selective measurement in complex media using a lifetime-based or a time-gated detection

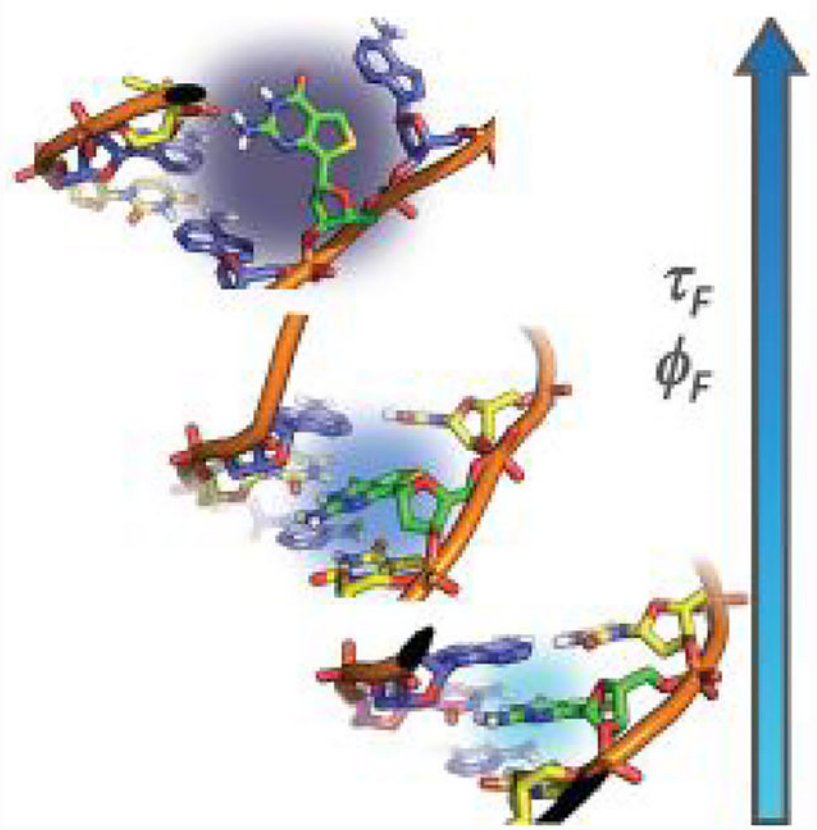
[*]Corresponding author: yves.mely@unistra.fr.
[+]These authors contributed equally to this work and were ordered alphabetically.

Supplementary Information

Additional computational details: Building the computational model, Spectroscopic properties of thG-labelled (−)/(+)PBS duplexes, Additional results on the FC region, Additional results on the S1minima, Spectroscopic properties and molecular dynamics simulation of thG-labelled (−)/(+)PBS mismatched duplexes; Additional analysis of the effects modulating thG photophysics in DNA.

scheme. Taken together, our data highlights thG as an outstanding emissive substitute for G with good QY, long fluorescence lifetimes and exquisite sensitivity to local structural changes.

Graphical Abstract



1. Introduction

DNA molecules are fundamental actors in living organisms as they encode genetic information needed for development, function and reproduction. Such processes rely on interactions with a large set of proteins that can impact the structure, conformation, bending or dynamics of the DNA double helix. These effects can be localized, as in base-flipping and bending events, or affect the DNA molecules globally, when compacted by histones or unwound by helicases or nucleic acid chaperones. To interrogate the molecular mechanisms that govern such protein-mediated variations, fluorescence techniques are a tool of choice due to their high sensitivity and exquisite spatio-temporal resolution. However, as the natural nucleobases are essentially non-fluorescent, there is a need for suitable emissive nucleobase analogues, which are able to seamlessly substitute the natural nucleobases while displaying bright and environmentally sensitive emission. The recently developed thienoguanosine (thG), an isomorphic guanosine (G) surrogate, appears particularly promising in this respect, almost perfectly replacing the natural G in both matched and mismatched duplexes.[1,2] Moreover, unlike many other fluorescent nucleobase analogues, thG is not strongly quenched when incorporated into DNA oligonucleotides and can thus be used to selectively and

faithfully monitor the conformation and dynamics of a given G residue in a DNA sequence[3]. The only anticipated limitation of thG may be its imperfect ability to form higher DNA structures (i.e., triplexes, quadruplexes) due to its lack of basic nitrogen corresponding to N7 in the purine skeleton.

Importantly, thG exhibits rich information content, as it can exist in two tautomeric forms (i.e. H1 and H3 keto-amino tautomers) at neutral pH, exhibiting considerably shifted absorption and emission spectra[4] and different fluorescent lifetimes[5,6]. The equilibrium between these tautomers informs on the local environment of thG. Due to its unique properties, thG was successfully used to monitor the local conformational changes associated to the binding of the HIV-1 nucleocapsid protein NCp7 to the (−)Primer Binding Site (PBS) and to decipher the mechanism of the NCp7-promoted (+)/(−)PBS annealing reaction[7]. The fluorescent nucleoside analogue in its deoxyribo or in its ribo form has also been successfully used to monitor B to Z transitions of DNA[2], single nucleotide polymorphism[7], distance measurements in DNA[8,9], base flipping reactions[10], ribozyme-mediated cleavage[11], and cellular activity of siRNAs[12].

Most of the applications of thG reported so far rely on empirical changes in fluorescence intensity and thus, do not exploit the full potential of this information-rich probe. The main reason is that a full understanding of thG photophysical properties, especially when thG is included in DNA, is still missing. Such an understanding would be useful for extracting additional information about any conformational changes or protein binding events by interpreting the observed changes in the photophysics on the basis of structural and electronic effects. For instance, this knowledge would be useful to explain the previously observed 2.5-fold difference in the quantum yield (QY) and mean fluorescence lifetime between corresponding matched and mismatched thG-labeled duplexes[3]. This large change in the fluorescence properties highlights the exquisite sensitivity of thG to its environment. Its understanding would be critical for gaining insight into the influencing factors. Such knowledge would also be very useful for laying the foundations for future applications of this probe and, in particular, for selecting the appropriate sequences for optimal spectroscopic changes in the presence of a given protein or other external effectors and ligands.

To augment the understanding of the photophysics of thG-labeled duplexes and, notably, their dependence on the DNA sequence context, we investigate herein the steady-state and time-resolved fluorescence properties of thG-labeled DNA duplexes, where we systematically vary the composition of neighboring and opposite bases of thG. As a model DNA duplex, we use the 18-mer (+)/(−)PBS duplex in which the G7 residue of (−)PBS was replaced by thG (Figure 1). The experimental data were complemented by molecular dynamics (MD) simulations to explore the structural behaviour of the investigated duplexes, as well as by mixed quantum mechanical/molecular mechanics (QM/MM) calculations to map the main photophysical processes in each duplex. Taken together, our integrated analysis provides a rationale for the observed spectroscopic properties and their dependence on thG's flanking or opposite nucleobases, disentangling the different factors modulating the photophysical features of thG.

2 Materials and Methods

(−)PBS DNA sequences labelled by thienoguanosine (thG) at 7th position (Figure 1) were purchased from TriLink Biotechnologies (USA). The complementary matched and mismatched (+)PBS DNA sequences were purchased from IBA Nucleic Acids Product Supply (Germany). Note, a stable tetrahydrofuran residue was used as a substitute for a "true" apurinic site. This commonly used substitution has been demonstrated to be a structurally and functionally viable surrogate[13–17]. thG-labeled (−)/(+)PBS duplexes were obtained by heating 10 μM thG7(−)PBS and 30 μM (+)PBS at 85 °C for 3 min and then, slowly cooling them down to room temperature. All experiments were done in 25 mM TRIS-HCl buffer pH=7.5, 30 mM NaCl and 0.2 mM $MgCl_2$.

Absorption and fluorescence measurements

Absorption spectra were collected on a Cary 4000 UV-visible spectrophotometer (Varian). Fluorescence spectra were recorded at 20°C on a FluoroLog spectrofluorometer (Jobin Yvon) equipped with a thermostated cell compartment. Matched and mismatched thG-labeled duplexes were excited at 315 nm and 360 nm wavelengths for steady-state fluorescence studies. The spectra were corrected for buffer fluorescence, lamp fluctuations, and detector spectral sensitivity. Quantum yields (QYs) of the thG-labeled duplexes were determined using quinine sulphate (QY=0.546 in 0.5 M H_2SO_4) as a reference[18]. Time-resolved fluorescence measurements were performed with the time-correlated single-photon counting technique[19]. Excitation wavelength at 360 nm was generated by a supercontinuum laser (NKT Photonics SuperK Extreme) with 10 MHz repetition rate. The fluorescence emission was collected at different wavelengths across the spectrum through a polarizer set at the magic angle and a 16 mm bandpass monochromator (Jobin Yvon). The single-photon events were detected with a micro-channel plate photomultiplier R3809U Hamamatsu coupled to a pulse pre-amplifier HFAC (Becker-Hickl GmbH) and recorded on a time-correlated single photon counting board SPC-130 (Becker-Hickl GmbH). The instrumental response function (IRF) recorded with a polished aluminium reflector was fitted with a Gaussian function (FWHM=50 ps). The measured fluorescence decays were first deconvoluted with the instrumental response function and fitted using the maximum entropy method[20,21] to assess the number of fluorescence lifetimes in the decays. Then, time-resolved exponential decays were fitted by a global analysis algorithm. The fitting function was a multi-exponential decay convolved with a normalized Gaussian curve of standard deviation σ standing for the temporal IRF and a Heavyside function. The fitting function was built in Igor Pro (Wavemetrics). All emission decays were fitted using a weighting that corresponds to the standard deviation of the photon number squared root. The lifetimes were shared for all emission wavelengths while the amplitudes were kept as free parameters during the fitting (Levenberg–Marquardt algorithm). A very good match was observed for the lifetime values obtained by the maximum entropy method and the global analysis. The results of the global fitting were used to construct the decay associated spectra (DAS) using: $I_i(\lambda)=I(\lambda)\alpha_i(\lambda)\tau_i/\Sigma\alpha_i(\lambda)\tau_i$, where $I(\lambda)$ is the steady-state emission spectrum and $\alpha_i(\lambda)$ are the wavelength-dependent amplitudes. The DAS provide the spectra associated to each lifetime component, and thus indicate the contribution in intensity of each species to the whole emission of the duplex. The real amplitudes associated to each lifetime component

were calculated from: $A_i = \frac{\int I_i(\lambda)d\lambda}{\int I(\lambda)d\lambda}$, where the numerator is the area under the DAS associated to the species i and the denominator is the area under the steady-state emission spectrum.

Melting temperatures

Melting curves were recorded by monitoring the changes in the absorbance at 260 nm of solutions containing final concentrations of 2 μM (−)PBS and 2.2 μM (+)PBS in in 25 mM TRIS-HCl buffer pH=7.5, 30 mM NaCl and 0.2 mM $MgCl_2$. The two sequences were annealed by a heating/cooling cycle prior to the measurement of the thermal denaturation. The melting curves were recorded in a 40–90°C range on a Cary 4000 spectrophotometer (Varian/Agilent) equipped with a Peltier thermostatted cell holder. The heating speed was 1°C/min. 200 μL of mineral oil were added on the top of the DNA solutions to avoid evaporation. Melting curves were converted into a plot of α versus temperature, where α represents the fraction of single strands. The melting temperatures were extracted from these curves by calculating their first derivative.

The melting temperatures of the different duplexes in their unlabelled form were also predicted by using the DINAMelt web server (http://unafold.rna.albany.edu/). The concentrations of PBS sequences and salt concentrations used for the predictions corresponded to the experimental ones.

Molecular dynamics simulations

DNA duplexes were built by the nucleic acid builder (NAB) molecular manipulation language[22], while parameters for thG and other MD settings were retrieved from a previous work[4]. Briefly, MD simulations were performed with AMBER 16[23], using the OL15 force field for DNA and the GAFF2 force field for thG. A box of TIP3P type water molecules buffering 8 Å from the solute's surface was built, while Na^+ ions were added to achieve the system neutrality. Analysis of MD trajectories was carried out by cpptraj[24], while the calculation of the delta energy of binding between the two strands in the duplex was performed with the MMPBSA.py script[25]. To check the conformational and thermodynamic stability of matched and mismatched (+)/(−)PBS duplexes, MD simulations were carried out on the most representative systems. All the systems were submitted to 200 ns of unrestrained MD production in explicit solvent, while analysis was carried out at the conformational (cluster analysis) and thermodynamic (calculation of the delta energy of binding) levels. Two replicas of MD simulations were carried out for all matched and mismatched duplexes. For systems showing flipped out thG residues, a third replica was carried out as well. The main atoms forming the nucleobase cycles were considered in the definition of the center of mass, which was used in the measurements of base-base distances along MD trajectories.

QM/MM calculations

Computational details: The ground- and excited-state minima of all the duplexes have been optimized by QM/MM calculations, using the ONIOM[26] interface as implemented in Gaussian09[27]. Density Functional Theory (DFT) and its time-dependent version (TD-DFT), adopting M052X[28,29] functional and 6–31G(d) basis set, was selected as QM method. This

functional was shown to well reproduce the tautomeric equilibrium in thG and the related absorption spectra, with the exception of a small blue shift (0.2–0.3 eV)[5]. Furthermore, M052X can reliably treat dispersion interactions and electronic transitions with Charge Transfer (CT) character[29], making it suitable for the study of a strongly coupled multi-chromophore system as DNA, as shown by many previous studies[30]. We chose the Amber parm96.dat Force Field for MM calculations[31]. Bulk solvent (water) effect has been included (in both QM and MM parts) by the Polarizable Continuum Model (PCM)[32,33]. Ground- and excited-state minima have been optimized by using the Linear Response version of the PCM, for which analytic gradients are available. We instead resorted to State Specific (SS-PCM)/TD-DFT calculations to provide a more accurate estimate of electronic transitions with significant CT character[34,35]. Further analysis concerning excited-state densities, molecular charges, and energetics were done considering a simplified model including only the four bases which were considered at the QM level and including environmental effects by PCM. In this latter case, however, we have used the dielectric constant of $CHCl_3$, 4.8, which can better mimic that experienced by a base within a duplex[36,37].

Computational models: Our computational analysis has been performed on a model including 8 bases (Figure S1), with the phospho-deoxyribose backbone and Na^+ counterions, namely bases 5, 6, 7 (thG), 8 and the corresponding bases 5’−8’ of the complementary strand. In our computations, four bases have been treated at the QM level, whereas the remaining ones, together with the phospho-deoxyribose backbone and Na^+ counterions, were described at the MM level. For both matched and mismatched systems, the QM region includes thG at position 7 and its Watson-Crick partner as well as the stacked base pair closest to thG, according to MD simulations. A stacked dimer appears the minimal model that, though being still computationally tractable, is able to provide a qualitative correct understanding of DNA photoactivated dynamics. Actually, there are convergent experimental and computational indications that, though the spectroscopic states in absorption can be delocalized over multiple bases, the main photophysical and photochemical processes in DNA involve one base or two stacked bases[38,39]. Nevertheless, the choice of the ‘stacked’ partner of thG to be included in the QM part is an additional source of ambiguity. To address this question, two models were considered for the matched TthGT/ACA duplex where thG is flanked by T residues. In model 1, T6 and A6’ (in addition to thG7 and C7’) are included in the QM region whereas T8 and A8’ are in the MM region. On the opposite, in model 2, T8 and A8’ are instead included in the QM region whereas T6 and A6’ are in the MM region. In this last case, the model has been extended to include also bases 9 and 9’. Since T6 and T8 are not symmetrically stacked to thG, in this way, we can get insights on the effect of conformational heterogeneity on the spectral properties of thG.

Starting from the most representative clusters issuing from MD simulations, we have optimized the ground state QM/MM minima for all models. In a second step, we have characterized the lower energy excited states in the Franck-Condon (FC) region and their Potential Energy Surfaces at the PCM/TD-M052X/6–31G(d)/MM level. Our re-optimization procedure, necessary in order to avoid any artefact due to the use of cluster average geometry and the MM field in the MD simulations, is also prone to predict a too close

stacking geometry between thG and the adjacent stacked residue treated at the QM level. Two limitations are the main responsible of this result. First, the other base stacked to thG is treated at the MM level, favouring a closer approach to that described at the QM level. Second, the effects of conformational fluctuations of the DNA are underestimated (a single minimum is analysed).

3. Results

To explore the photophysics of thG in DNA oligonucleotides and its dependence on sequence context, the 18-mer (+)/(−)PBS duplex was used, where the G7 residue in (−)PBS was replaced with the isomorphic emissive nucleoside thG, and the flanking T6 and T8 residues were systematically replaced by A, C or G residues (Figure 1). The complementary residues on (+)PBS were changed accordingly to maintain Watson-Crick pairing. We also explored the impact of the base opposite to thG, by replacing its complementary C residue in (+)PBS by A, T, G or an abasic residue, thus intentionally generating a mismatch. In total, the absorption and emission properties of 20 (+)/(−)PBS duplexes, including 4 fully matched duplexes and 16 mismatched duplexes, were analysed.

3.1 Matched Sequences

3.1.1. Spectral characterization—The absorption spectra of all matched duplexes (with C opposite to thG) were recorded (Figure 2A). As the main absorption band of thG is red-shifted compared to the natural bases, it could be compared to the spectrum of thG H1 tautomer, which has previously been shown to be the only tautomer present in matched (+)/(−)PBS duplexes[4]. A slight red-shift and a large decrease in the extinction coefficient was observed for thG in matched duplexes, as compared to the free thG H1 tautomer (Figure 2A, green spectrum). This strong hypochromic effect (29–41%, Table 1) clearly suggests that Watson-Crick pairing with a complementary C residue enables optimal stacking of thG with the flanking nucleobases[40], especially when the thG–C base pair was flanked with another G–C base pairs (Figure 2A, blue and orange spectra).

Emission spectra were recorded upon excitation at 360 nm (where only the H1 tautomer absorbs) and 315 nm (where both the H1 and H3 tautomers absorb) (Figure S2). When normalized, the emission spectra of all matched duplexes perfectly overlapped at the two excitation wavelengths, confirming that the H3 tautomer was essentially absent in these duplexes[4]. The emission maxima of the duplexes are blue-shifted by ~10 nm (457–463 nm, Table S1) as compared to the free thG H1 tautomer (470 nm) in water[4] and are thus comparable to that the free thG H1 in methanol (458 nm). Interestingly, the duplexes exhibited rather low QYs (0.15–0.18, Table 1) almost independent on the nature of the flanking residues, in agreement with previous data on a 12 bp duplex, where thG at different positions (and thus, with different neighbors) was found to exhibit similar QY values[10]. The emission spectra of the matched duplexes were also checked to be independent on the NaCl concentration in the 30 – 150 mM range (Fig. S3A).

Next, time-resolved fluorescence decays were recorded for the thG-labelled matched duplexes over the entire emission spectrum of thG (420–560 nm) to recover the decay associated spectra (DAS) (Figure 3). For all matched duplexes, the thG fluorescence decay

was well fitted with two lifetime components (Figure S4A). The long-lived lifetime τ_2 (10.6–12.3 ns) was almost independent of the neighboring nucleobases' identity (Table 1), in line with the observed QYs. The short-lived component τ_1, displayed greater variations, with lower values being associated with neighboring A and T (1.7–2.4 ns), as compared to C and G (3.4–4.4 ns). With respect to amplitudes, the long-lived lifetime, τ_2, was the dominant species in all duplexes, but its contribution was more pronounced with A and T (83–86%) than with C and G (61–69%) as neighboring bases. As a result, the DAS associated with the long-lived lifetime nearly overlapped the steady-state spectrum when thG was flanked with A or T (Figure 3A and B), while the short-lived lifetime significantly contributed to the emission of the duplexes when thG was flanked with G or C (Figure 3C and D). From the comparison of the QY and mean excited-state lifetimes, the radiative, k_r, and non-radiative, k_{nr}, rate constants could be calculated. The k_r values (1.5–1.9×10^7 s^{-1}) of thG in the matched duplexes were significantly lower than those of the thG H1 monomer in water (2.5×10^7 s^{-1}) or methanol (2.9×10^7 s^{-1}). As the k_r values are proportional to the extinction coefficients of the chromophores[41], the decreased k_r values are in line with the hypochromism observed in the absorption spectra (Figure 2A and Table 1). Moreover, our data further revealed that the lower QY values of thG in the duplexes as compared to free thG H1 in water were mainly due to an about four-fold increase in the k_{nr} value, suggesting the presence of additional non radiative pathways as compared to the thG H1 monomeric nucleoside.

3.1.2 MD Analysis—MD simulations on matched duplexes indicated that replacement of the T bases flanking thG with A, C, or G had a minimal structural effect, since the global conformation of the duplex as well as the thG-C Watson-Crick base pair was preserved (Figures 4 and S5). From the cluster analysis of MD simulations, a single representative cluster containing more than 99% of MD frames was extracted for each duplex.

The thermodynamics stability of the matched thG-labelled duplexes was estimated by the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) approach by calculating the delta energy of binding, ΔEb (Table 2), which corresponds to the difference in total energy between the duplexes and their constitutive single-stranded sequences and is thus related to the hybridization enthalpy.[42] The duplex stability was also characterized by its melting temperature T_m, either determined experimentally with the thG-labelled duplexes from the temperature dependence of their absorbance changes at 260 nm (Figure S6) or predicted by DINAMelt[43] on the corresponding unlabelled duplexes (Table 2). Interestingly, an excellent match was found between the experimental and the DINAMelt-predicted *Tm* values, confirming that substitution of G by thG does not impact the melting temperature in DNA duplexes.[2,10] The only deviations between experimental and calculated *Tm* were observed for the duplexes with abasic sites, probably due to the fact that DINAMelt is not optimized for abasic sites.

As expected from the higher stability of G-C base pairs as compared to A-T base pairs, the GthGG/CCC and CthGC/GCG duplexes were more stable by 5–6°C as compared to the AthGA/TCT and TthGT/ACA duplexes (Table 2).

3.1.3. QM/MM calculations—After having optimized the minima for the matched sequences (Figures 5A and S1) at the QM(PCM/M052X/6–31G(d))/MM level, we have

computed the 12 lowest energy transitions at the TD-M052X levels (Tables 3 and S2), in order to be sure to include all the main electronic transitions involving thG, including the CT ones.

When compared to the free thG H1 in water, the predicted moderate red shift of the absorption maximum and strong hypochromicity (~30%) of the transition (Table 3) are in line with the experimental data (Figure 2A and Table 1). Inspection of the lowest energy transitions (Table S2) shows that in addition to the spectroscopic $\pi\pi^*$ state, essentially localized on thG, a relatively close-lying CT transition (<1 eV less stable than S_1), involving thG and a flanking base is always present. However, its direction depends on the nature of the neighbour base, since thG serves as electron donor when stacked with C but an electron acceptor when stacked with G (Figure 5B).

In the second step, we have optimized both S_1 and the lowest energy CT transition at the QM/MM level for all matched systems. The main features of the resulting minima are summarized in Table 4.

The high stability of the CT minima strongly suggests that an additional, non-emissive, deactivation channel is available for thG within duplexes, in line with the multi-exponential fluorescence decay and the decrease of both fluorescence QY and lifetimes observed experimentally when thG is included in DNA. The comparison between the different duplexes indicates that the CT minimum is relatively less stable for the (thG→A) CT state (Table S3), in line with the slightly higher experimental QY for the matched AthGA/TCT duplex (Table 1). Only for the GthGG/CCC duplex, thG acts as an electron acceptor in CT, whereas thG serves as electron donor for the other systems (Table 4). Our calculations also show that the minimum of the $\pi(^{th}G)\rightarrow \pi^*(^{th}G)$ excited state acquires a small CT character ($0.01 \leq {}^{LR}CT \leq 0.03$), since some contribution of the stacked bases is found. This leads to a decrease of the oscillator strength and thus, of the emission intensity as well as an increase of the radiative lifetime with respect to thG H1 tautomer in water, in line with the experiments (Table 1). Moreover, using the experimental QY values, the mean lifetime values (last column of Table 4) recalculated by the product of QY with the calculated k_r value match very well with the experimental values (Table 1). Comparison of the results obtained for the two models used to study the TthGT/ACA duplex provides interesting indications on the effect of the variation of the stacking geometry (due for example to conformational fluctuations of the duplex) on the thG optical properties. As explained in the Materials and Methods section, the two models differ by the T base (5' or 3' to thG) included in the QM region. Model 2 monitors the effect of T8 at the 3' end with respect to thG. MD simulations and QM/MM calculations show that T8 is strongly stacked to thG (with a closer stacking distance and a larger overlap between the rings), leading to a minimum with strong $^{th}G^{+}\rightarrow T^{-}$ CT character. In Model 1, where the focus is on T6, our calculations show that T6 is more loosely stacked with thG, leading to an excited-state minimum localized on thG. This clearly indicates that the photophysics of thG is primarily governed by the interaction with its nearest flanking base which most strongly interacts with thG.

3.2. Mismatched duplexes

For all mismatched duplexes, comparison of the spectra at 315 nm and 360 nm excitation wavelengths revealed an only marginal contribution of the H3 tautomer (Figure S7). Its contribution was most prominent with opposite A, but represented less than 10%. Due to its marginal contribution, no accurate analysis of the H3 tautomer could be performed. Therefore, all the measurements in the remaining part of this study were performed at 360 nm, where only the H1 tautomer absorbs. As for the matched duplexes, the spectra of the mismatched duplexes were found to be independent on the NaCl concentration in the 30–150 mM range (Figure S3B and C). For all duplexes, the fluorescence decays were biexponential over the whole spectral range (Table 5, Figure S4B and C). The short-lived lifetimes τ_1 were rather homogeneous (2–5.5 ns) and contributed to less than 10% of the total emission of thG (as shown by their f_1 values) in 12 out of the 16 tested duplexes. As a consequence, the emission of the mismatched duplexes was largely dominated by the long-lived lifetimes τ_2, which varied in a large range (9.3 – 28.6 ns) and were therefore mainly discussed in the following. Interestingly, the k_r values were distributed in a narrow range (1.4–2.1×10^7 s^{-1}) indicating that the QY and $<\tau>$ values change in a highly correlated manner and thus, that dark species, with lifetimes under the 20 ps resolution of our set-up, are likely marginal in all duplexes.

3.2.1 Mismatched duplexes with flanking G and C bases—In mismatched duplexes where thG is flanked by G residues, the absorption spectra show a strong hypochromism (up to 37%), except for the GthGG/CAC duplex where thG is mismatched with A (13%) (Table 5). Among the mismatched duplexes, those with flanking G residues exhibit the lowest QY, τ_2 and $<\tau>$ values, being comparable to the values of the corresponding matched duplex (Table 1). The only exception is for the GthGG/CGC duplex, showing an only two-fold decrease of its QY as compared to free thG H1 in water, and a τ_2 value comparable to the thG H1 lifetime in water. Another striking feature is the high amplitude (~ 50%) associated to the short-lived lifetime τ_1 of the GthGG/CAbC duplex with an abasic site, so that this lifetime strongly contributes to the emission spectra (Figure S8).

The spectroscopic properties of mismatched duplexes with thG flanked by C bases are not too different from those with flanking G bases. The observed hypochromism is fairly high (24–30%), with again the exception of the CthGC/GAG duplex with thG opposite A (15%). The QY, τ_2 and $<\tau>$ values are comparable to those observed when thG is flanked by G, with the exception of the thG-A mismatch, where the QY, τ_2 and $<\tau>$ values are 1.5 to 2-fold higher for the duplex with flanking C residues. Finally, the short-lived fluorescence lifetime of the duplex with an abasic base was found to exhibit a very high amplitude (68%) and thus, strongly contributes to the overall emission (Figure S9).

To further understand the spectroscopic data of the mismatched duplexes with thG flanked by G or C bases, we performed MD and QM/MM calculations. The GthGG/CTC and CthGC/GTG duplexes with a thG-T mismatch show rather high Tm values (68–69 °C) and stable ΔEb values (−66 kcal/mol), as a result of the Wobble-like conformation adopted by the thG-T mismatch (Figure 6A–B). This conformation is stable over time as shown by the RMSD analysis (Figure S10A) and the observation of a single cluster of MD frames with a

population over 99%, which is structurally close to the matched duplex (Figure S10B). The similarity in structure and stability of these mismatched duplexes with the matched duplexes is well consistent with the similarity of their spectroscopic properties.

The GthGG/CGC and CthGC/GGG duplexes with thG-G mismatch show lower stability, with Tm ~ 67°C and ΔEb ~ −60 kcal/mol. In contrast to the thG-T mismatch, thG is unable to establish a Wobble-like pairing with an opposite G. The amino group of thG forms an H-bond with the carbonyl group of G, while the remaining portion of thG is exposed towards the solvent (Figure 6C–D). By monitoring the distance between the amino group of thG and the carbonyl group of opposite G, thG is partially flipped out from the duplex for a short period that corresponds to 13% of MD trajectory (Figure S11A). Both the solvent exposure of thG and its ability to flip out are well consistent with the τ_2 lifetime value close to that of the free nucleobase in water.

The GthGG/CAC duplexes show a high stability (Tm = 68–69 °C and ΔEb = −73 kcal/mol), likely due to the ability of thG to establish H-bonds with the opposite A (Figure 6E–F) and properly stack with the flanking G bases. As a result, thG remains within the duplex for the majority of MD simulation time and its photophysical properties do not deviate too much from those of the corresponding matched duplex. In contrast, there is an additional destabilization of 7.4 kcal/mol and a 4 °C decrease in T_m for the CthGC/GAG duplex. This additional destabilization likely reduces the base stacking of thG with its flanking C bases, explaining the substantial increase in the QY, $<\tau>$ and τ_2 values.

Simulation of the duplexes with an abasic site opposite to thG using the parameters previously reported by Chen et al.[44] indicates that the abasic site does not dramatically affect the conformation of thG, which is stacked within the duplex by flanking bases and eventually connected to the opposite sugar by water-bridged H-bond interactions (Figure S12A and B). For the GthGG/CAbC duplex, the distance between the centers of mass of the main rings of thG and G6, the closest flanking nucleobase along the MD trajectory was found remarkably constant and similar to the corresponding distance in the matched GthGG/CCC duplex (Figure S13A), in agreement with the highly similar photophysical properties of thG in the two duplexes. Noteworthy, the distance between thG and G8 has larger fluctuations and is on the average longer in the mismatched duplex than in the matched duplex. Nevertheless, in a few fluctuations, G8 came very close to thG in the mismatched duplex. Interestingly, the mismatched CthGC/GAbG duplex with flanking C residues was less stable by 1.5 °C than the GthGG/CAbC duplex with flanking G residues (Table 2). This lower stability is also perceived in the highly fluctuating distances of thG with its C6 and C8 neighbors (Figure S13B). These fluctuations, notably with the closest C6 residue, induce an overall increase in the distance with thG as compared to the CthGC/GCG matched duplex, in line with the increase in the τ_2 value observed in the mismatched duplex (compare Tables 5 and 1). Interestingly, the distance fluctuations between C8 and thG were very large, occasionally leading to very close proximity.

QM/MM calculations on the GthGG/CTC duplexes (Figure S1) confirm a strong interaction of thG with the close-lying G6 base, leading to a strong hypochromism (Table 3). Excited-state geometry optimizations of S_1 lead to a minimum with a very strong G→thG CT

character (Table 4), suggesting an efficient decay channel through this CT, in line with the low QY (0.10) of these duplexes. The coupling between the bright thG $\pi\pi^*$ state and the CT state is so large that it has not been possible to locate a $\pi\pi^*$ minimum on thG. However, as explained in Materials and Methods, our procedure likely overestimates the interaction between the stacked bases considered at the QM level and underestimates the conformational flexibility of the duplexes. This limitation is expected to be more important for mismatched than for matched duplexes. As a consequence, there are probably structures for which the decay to an excited-state localized on thG is favoured.

3.2.2 Mismatched duplexes with flanking A bases—In mismatched duplexes, where thG is flanked by A nucleobases, the hypochromism with respect to free thG was substantial (28–33%), with the exception again of the AthGA/TAT duplex with a thG-A mismatch (12%). All QY values (0.37–0.48) were close to that (0.51) of free thG H1 tautomer, indicating that flanking A residues had the lowest quenching potency on thG in mismatched duplexes (Table 5). The values of both τ_1 and τ_2 fluorescence lifetimes were rather uniform for the different duplexes with flanking As. Strikingly, the τ_2 values (25–28.6 ns) were substantially higher than the lifetime of free thG H1 in water (20.5 ns). This may be mainly attributed to the decrease of k_r from $2.5\times10^7 s^{-1}$ for free thG H1 in water to ~$2\times10^7 s^{-1}$ for thG in the duplexes, while the k_{nr} values were comparable to those of thG H1 in water. The τ_1 lifetimes (2.3–2.6 ns) exhibit small amplitudes and thus, contribute only to a very small extent (≤ 4%) to the overall emission of these duplexes (Figure S14). Thus, the lifetime data suggest that only a small thG population is quenched by its A neighbors in these mismatched duplexes.

Among the mismatched duplexes, the lowest stability (Tm = 60–62.8 °C and ΔEb between −49 and −56 kcal/mol) was observed when thG is flanked by A nucleobases (Table 2). For the AthGA/TGT and AthGA/TAT duplexes, MD simulations show that thG mostly occupies local single-stranded regions, being paired with the opposite +1 base or temporarily exposed to the solvent. In both cases, stacking interactions of thG with flanking base pairs are completely lost (Figure 7A–B), in full line with the high QY, $<\tau>$ and τ_2 values observed with these duplexes. When thG was mismatched with T in the AthGA/TTT duplex, it establishes a Wobble-like base pair (Figure 7C), which marginally stabilizes the duplex (T_m = 60.4 °C). In line with this low stability, the distance between the centers of mass of the main rings of thG and the closest flanking adenine A6 along the MD trajectory is highly fluctuating (Figure S15A) leading to conformations where A6 is highly shifted in respect to thG (Figure S16B). This is in sharp contrast to the AthGA/TCT matched duplex, where this distance is consistent (Figure S15A), keeping thG and A6 in close proximity (Figure S16A). These simulations are in agreement with the high QY, $<\tau>$ and τ_2 values observed with the AthGA/TTT duplex. Low stability (T_m = 60°C) was observed also for the AthGA/TAbT duplex with abasic site (Figure S12C), although the relative position of thG with respect to the flanking bases does not significantly deviate from that measured in the matched duplex (Figure S13C). In this particular case, the explicit water molecules included in the MD simulation may artificially stabilize the duplex structure, which can obscure the fluctuations in distances between thG and its neighbors occurring in this less stable duplex.

QM/MM calculations on the AthGA/TAT duplex where thG is opposite A (Figure S1) confirmed that thG is highly exposed to the solvent and does not strongly interact with the close-lying A nucleobases. As a consequence, the computed hypochromism is very small (Table 3) and the S_1 geometry optimization provides a picture similar to that for free thG H1, i.e. only a $\pi\pi^*$ minimum is optimized, as it can be noted by the estimated radiative lifetime (Table 4).

3.2.3 Mismatched duplexes with flanking T bases—All mismatched sequences with flanking T bases exhibit a substantial hypochromism (23–31%, Table 5). The QY, τ_2 and $<\tau>$ values are in between the values observed for duplexes with flanking A residues and those with flanking G or C residues. The effect of flanking T bases on thG photophysics is therefore smaller than that of G or C residues and larger than that of A residues (Table 5). The τ_1 values are rather high (3.0–5.5 ns) but contribute only little (4–9%) to the overall emission of these duplexes (Figure S17 and Table 5). Of note, the data of the TthGT/ATA duplex with opposite T in Table 5 should be considered as a refinement of our data on the same duplex in a previous paper[3], which were very similar, except for the presence of small population of dark species. This small population was not found anymore and probably resulted from small errors in the determination of the QY.

As for the mismatched duplexes where thG is flanked by A nucleobases, low stability ($T_m \leq$ 63 °C and ΔEb between −40 and −59 kcal/mol) was observed for the mismatched duplexes with flanking T nucleobases. For the TthGT/ATA duplex with a thG-T mismatch, MD simulation show that thG interacts with the mismatched T mostly by a Wobble-like pairing that is persistent in MD trajectories for around 70% of the simulation time (Figure 8A). In this major conformation, thG loses partially its stacking interaction with one of the neighbouring bases. The remaining fraction of MD simulation time is represented by a destabilized conformation (Figure 8B) in which thG does not interact with the opposite and flanking bases. Similarly, when thG is mismatched with A in the TthGT/AAA duplex it remains within the duplex for around 22.6% of MD simulation time (Figure S11B), while it moves towards the solvent accessible space for 68.5% of MD time (Figure 8C) and flips out from the duplex for the remaining 8.9% of MD simulation time (Figure S11B). In the TthGT/AGA duplex with a thG-G mismatch, thG is unable to establish direct interactions with the opposite nucleobase, but remains in the solvent-accessible space within the duplex for most of the MD simulation time (Figure 8D). Finally, for the TthGT/AAbA duplex, we also observed a significant conformational instability of the duplex within the abasic site, such as underlined by the representative structure (Figure S12D) and the analysis of the distance between thG and the flanking T nucleobases with respect to the TthGT/ACA matched duplex (Figure S13D). Both the loss of stacking interaction as well as the partial or full exposure of thG to solvent are well consistent with the high QY, $<\tau>$ and τ_2 values observed with these duplexes

QM/MM calculations were then performed with the TthGT/ATA and TthGT/AAA duplexes. Based on the MD simulation data, we have analysed two families (groups 1 and 2) for the duplexes where thG is opposite T and three different conformational families (groups 1, 2, 3) for the duplexes where thG is opposite A (Figure S1). These structures differ on the degree of thG stacking (Figure S1). As detailed in the SI, both the mismatches with T and A present

geometries (groups 1 and 2) where the S_1 excitation is delocalized over the closest base (A/T). The higher the delocalization, the larger is the charge transfer (CT) and the lower is the radiative constant. In the mismatch with A, there is also a monomer-like group 3, where the excitation is localized on thG without any coupling to other bases. Finally, for the TthGT/AGA duplex where thG is mismatched with G, thG is flipped out into the solvent and does not interact with the surrounding bases (Figure S1) in agreement with MD results. Although no clear CT minimum was located in this case (Table 4), its population cannot be ruled out, since an additional T base (T8, not included in this model) is close enough to provide such kind of minimum, as we found for the TthGT/ACA matched duplex.

4. Discussion

Being able to faithfully report on the structure and dynamics of G residues in nucleic acids (NAs), thG, an environmentally sensitive isomorphic fluorescent analogue of guanosine, has emerged as a major tool for characterizing NA structural changes and interactions with proteins. To further evaluate its potency and better interpret the changes in its spectroscopic properties upon external perturbations, it is essential to fully understand the mechanisms underlying the photophysics of thG when incorporated into NAs. Toward this end, we combined experiment and theory to study (−)/(+)PBS DNA duplexes, where the flanking and opposite bases of a central thG in the (−)PBS sequence were systematically varied.

A strong hypochromism in thG's absorption spectra was observed for both matched and mismatched sequences (up to 41% and 37%, respectively; Figure 1, Table 1 and 5). As for the natural bases in DNA[40], this hypochromism is likely due to exciton coupling between thG and its flanking bases. The strongest hypochromism (41%) was observed in GthGG/CCC and CthGC/GCG matched duplexes, where thG was flanked by G-C base pairs. It was substantially lower (~30%) for the AthGA/TCT and TthGT/ACA matched duplexes with flanking A-T base pairs. This difference likely results from the stiffness brought about by the three consecutive G-C like pairs stabilized by three hydrogen bonds each. This rigidity likely governs the strength of the stacking interaction of thG with the nearby bases, and thus, the extent of hypochromicity. In line with this conclusion, a strong hypochromism (24–37%) was also observed with GthGG/CTC and CthGC/GTG duplexes where thG forms a wobble base pair with T, stabilized by two hydrogen bonds (Figures 6 and 8A). In contrast, the lowest hypochromicity was observed when thG was opposite A, as a result of a steric clash between the two residues, which shifts thG toward the solvent (Figure 7A). The absorption spectra computed at the QM/MM calculations are fully consistent with the experimental results, confirming the general reliability of our conclusions.

The QY values in DNA were found to be lower than the QY of free thG H1 in water (0.51) and to vary in a large range (0.1–0.48), indicating that the quenching of thG in the duplexes is highly context dependent. However, a very different picture emerged for matched and mismatched duplexes. In matched duplexes, the QY values as well as the τ_2 and $<\tau>$ values were rather low, and minimally dependent on the nature of the bases flanking thG. This minimal dependence of the photophysical features of thG on its flanking nucleobases also applies when it is flanked by nucleobases of different nature, as shown by our ab initio calculations (Figure S18A–C) and by the independence of thG's QY when placed in

different positions and therefore with different neighbors (GthGG, CthGC, GthGC or CthGT) in a matched duplex.[10] QM/MM calculations indicated that the matched duplexes were characterized by a relatively close-lying CT transition (<1 eV less stable than S_1) with a strong CT character, involving thG and the adjacent base. thG in the excited state serves as an electron acceptor when it is flanked by G, whereas it serves as an electron donor for the other bases (Table 4). This CT transition is likely responsible of an additional, non-emissive, deactivation channel in the matched duplexes that is consistent with the measured increase in the k_{nr} value, responsible of the decrease in QY and lifetime values. This efficient CT transition with all neighboring bases in matched duplexes is correlated with the high stability of these duplexes (Table 2) that places thG in the proper orientation and distance to the flanking bases. These conditions of relatively high stability that facilitate efficient CT transition are also met with the GthGG/CTC and CthGC/GTG mismatched duplexes where thG forms a Wobble base pair with T and is flanked by G-C base pairs (Table 2). In these duplexes, these three base pairs are stabilized by a set of eight H bonds, which probably preserve the same orientation and distance of thG with the flanking bases as in the matched duplexes, and thus, yield QY, τ_2 and $<\tau>$ values (Table 5) close to those of the matched duplexes (Table 1). Calculations also revealed that the minimum of the $\pi(^{th}G) \rightarrow *(^{th}G)$ excited state in matched duplexes acquires a low CT character (Table 4), which reduces its oscillator strength and radiative constant with respect to free thG H1, and thus leads to a further decrease in the QY and lifetime values.

In contrast to matched sequences, the values of QY, τ_2 and $<\tau>$ in mismatched duplexes are highly dependent on the nature of thG's flanking bases and to a lesser extent on its opposite base. Our experiments indicate that the impact of the flanking bases on thG fluorescence varies in the following order: G>C>T>A. The strong quenching potency of G residues is not surprising, as G is generally the strongest quencher among the natural bases for a large number of fluorescent nucleobases[45–48] and fluorophores[49], due to its high reduction potential and thus, high electron donating properties[49,50]. In line with our experiments, computational analysis suggests that both electronic and structural effects modulate the fluorescence of thG in DNA. QM calculations indicate that the strongest quenching effects are observed when thG is stacked with another base and involves a CT state (e.g. $^{th}G^{+} \rightarrow$Pyrimidine^{-} or $G^{+} \rightarrow {}^{th}G^{-}$), leading to an effective non-radiative decay channel (Table 4). Moreover, inspection of the FC region (e.g. energy difference between the lowest energy CT state and the bright state) and analysis of our excited-state geometry optimizations suggest that the coupling between the bright spectroscopic state localized on thG with the CT state is larger when the flanking bases are G or C, and much smaller for A. This stronger coupling suggests a more effective decay to CT states for the former two bases. The availability of this additional decay channel obviously depends on the proximity of thG to its flanking base, as shown by the excellent correlation between the photophysical properties of thG and the distance fluctuations between thG and its closest flanking nucleobase over MD time (Figures S13 and S15). Thus, the local stiffness of the duplex and, more generally, its conformational behaviour, also modulate thG spectroscopic properties. On this ground, flanking G-C base pairs increase the local stiffness of the duplex, due to their three hydrogen bonds, explaining why flanking C residues are more effective quenchers than flanking T residues.

Interestingly, the τ_2 values of all mismatched duplexes with flanking A residues and two out of the four mismatched duplexes with flanking T residues (23.8–28.6 ns) were much higher than the fluorescence lifetime of thG H1 in water (20.5 ns). As fluorescence lifetimes are dictated by $1/(k_r+k_{nr})$, the large τ_2 values can be explained in part by their k_r values, which are lower ($1.7–2.2\times10^7$ s^{-1}) than that (2.5×10^7 s^{-1}) of thG H1 in water. The real k_r values associated to τ_2 may even be lower, since the k_r values reported in Table 5 were calculated from the mean lifetime, assuming that the k_r values are the same for both τ_1 and τ_2 lifetimes. As the k_r values are proportional to the chromophore's extinction coefficients,[51] the decreased k_r values can be related to the hypochromism observed for thG in all duplexes (Table 1). For all other duplexes, the observed hypochromism also reduces the k_r value in comparison to that of the thG H1 monomer, but its effect on the lifetime is masked by the large increase in k_{nr} due to the CT pathway in these duplexes.

Short-lived τ_1 lifetimes are observed for all the matched and mismatched duplexes, but their amplitudes are systematically higher when thG is flanked by G-C as compared to A-T base pairs. The amplitude and thus, the contribution of the τ_1 component to the overall emission is especially high in the highly dynamic context provided by an abasic site, where thG can be efficiently quenched by its G or C neighbors. This view is supported by MD simulations on mismatched duplexes with abasic site, showing that the distance between thG and its flanking G8 or C8 residues is highly fluctuating, occasionally leading to very short distances favourable for quenching (Figure S13A and B). These data suggest that the τ_1 lifetime might be related to a dynamic quenching of thG by CT to its neighbors, modulated by the conformational fluctuations of DNA.

From the complementary base perspective, the poor quenching of thG by its flanking bases in mismatches with A or G (Table 5) can be easily explained by the steric clash of thG with its opposite purine base, which dislodges the thG toward the solvent and prevents its stacking with the flanking bases (Figure 7). In contrast, a close contact and thus a strong quenching of thG by flanking C-G residues is observed when thG is opposite to a pyrimidine or an abasic site. This can be illustrated by the CthGC/GTG mismatched duplex where thG faces T and is flanked by C residues (Figure 6A and B). In this case, the most representative cluster shows that thG is close to both C and G residues on the opposite strand, in line with the observed low QY (0.15).

Overall, our experimental and computational results can be rationalized by a simple model, where the key interactions are those established between thG and its most closely stacked neighbours, which depending on the global conformational behaviour of the particular duplex examined, can be one of its flanking residues or less frequently, a base on the opposite strand. Increasing numbers of experimental and computational contributions indicate that the 'stacked dimer' is the most meaningful minima model to understand the photophysics and photochemistry of nucleic acids.[30,52,53] On this ground, we can explain why, as discussed in detail in section 6 of the SI, other effects, as those related to the presence of two different flanking residues for thG (Figure S18A–C), play a minor role in its photophysics (Table S4 and Figure S19). Analogously, the nature of the next closest neighbour, even in the extreme case of a GG step (Figure S18D), does not have a significant direct effect on the CT excited states in which thG is involved (Table S5). On the other hand,

it is clear that any change in the sequence, especially for mismatched duplexes can have a substantial and not a priori predictable effect on the photophysics of thG, but this study highlights the major chemical physical effects that have to be taken into consideration.

5. Conclusions

By combining steady-state and time-resolved fluorescence spectroscopy with molecular modelling and quantum mechanical calculations, the photophysical properties of thG in matched and mismatched DNA duplexes were collected, interpreted and rationalized. In all twenty sequences examined, thG was significantly emissive (QY 0.10–0.48) and exhibited a simple biexponential fluorescence decay, confirming its utility as a highly sensitive and robust probe in DNA duplexes. Due to its high QY, the brightness ($\varepsilon \times$QY) of thG in DNA duplexes is between 500 and 2250 M^{-1} cm^{-1} and thus, not too distinct from that of the brightest fluorescent nucleoside ever reported in nucleic acids (2600 M^{-1} cm^{-1}).[54] Moreover, the long-lived excited-state lifetime τ_2, which contributes to more than 75% of the total emission for 19 out of 20 of the tested duplexes (Table 5), shows unusually high values (9–29 ns). Such long-lived excited states are rarely observed with common small fluorophores, potentially facilitating selective measurements even in complex mixtures of fluorescently labelled molecules or in autofluorescent media by using a fluorescence lifetime-based or time-gated detection scheme.

For matched duplexes, the emission quantum yields (0.15–0.18), as well as τ_2 (10–12 ns) and $<\tau>$ (8–10 ns) were found to narrowly vary. The minimal dependence of the QY and excited-state lifetime values on thG's neighbors in matched duplexes appears to be advantageous, because thG can be used to replace any G residue in a DNA duplex, while maintaining similar photophysical properties. In this respect, the values given here can be considered representative for referencing in future studies.

In sharp contrast, the QY, τ_2 and $<\tau>$ values of thG in mismatched duplexes strongly depend on the nature of its flanking and paired residues. An obvious application for thG-labelled sequences emerging from our observations is the detection of single nucleotide polymorphism. This application was already suggested for thG on the basis of the analysis of the limited but context dependent contribution of the thG-H3 tautomer, when the duplexes were excited at 310 nm.[4] Herein, comparison of Table 1 and 5 shows clear differences in the QY, τ_2 and $<\tau>$ values when the matched C is replaced by any other nucleotide or an abasic site. Thus, the QY, τ_2 and $<\tau>$ values appear to reflect a simpler and more sensitive approach to detect single nucleotide polymorphism and identify the nature of the mismatched nucleotide. Moreover, comparison of the matched with mismatched duplexes further indicates that the context insensitive QY, τ_2 and $<\tau>$ values strictly depend on the local stability and stiffness of the matched duplexes, which ensure the strength of the stacking interactions of thG with the nearby bases. The local destabilization and structural changes in DNA induced by a mismatch or an abasic site opposite to thG are sufficient to restore a strong dependence of thG photophysics on its context, and notably on the nature of its flanking bases. In contrast to matched duplexes, the photophysical properties of thG flanked by nucleobases of different nature are more difficult to predict in mismatched duplexes, as they obviously depend on the exact geometry adopted by thG and its flanking

nucleobases as a result of the local destabilization induced by the mismatch. To our knowledge, thG's exquisite photophysical sensitivity to its context, while preserving good QYs, is unique among fluorescent nucleoside analogues[46,55] and is thus ideal for monitoring local structural changes induced by DNA ligands acting as helix destabilizers, NA chaperones, DNA bending proteins or base flipping proteins. This conclusion is fully in line with previous observations and helps explaining the strong changes in thG fluorescence observed when thG-labelled DNA duplexes or stem-loops interacted with the base flipping domain of UHRF1[10] or the HIV-1 nucleocapsid protein NCp7, acting as a NA chaperone.[7] In both cases, the protein-induced base flipping prevents the stacking between thG and its flanking nucleobases, which impedes the CT mechanism and thus, explains the observed fluorescence increase. Similarly, the fluorescence increase observed for a ribozyme-mediated cleavage that releases a short duplex with thG at its 5' terminus[11] could be easily explained by a dramatic decrease in CT efficiency as a result of the loss of one of thG's flanking residues and the poor stability of the duplex termini.

Taken together, our integrated analysis provides a rationale for the observed spectroscopic properties of thG in DNA duplexes and enables disentangling the different effects modulating its photophysical features. This analysis validates thG as a very unique fluorescent nucleoside surrogate that combines almost perfect substitution of G residues, preservation of good QY, unique long-lived excited-state lifetimes and exquisite sensitivity to solvent exposure and DNA context. This study opens new possibilities for the rational use and proper data interpretation of thG-labelled oligonucleotides in applications such as single nucleotide polymorphism and DNA/protein interactions. Applications of thG photoexcited dynamics, whose assessment is fundamental for the study of photoinduced oxidative damage, can also be envisioned.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank R. Sharma for performing preliminary experiments of this project. This work was supported by the Agence Nationale de la Recherche (ANR blanc PICO2 and SMFLUONA), the Labex NIE and the Centre National pour la Recherche Scientifique (CNRS). Y. M. is grateful to the Institut Universitaire de France (IUF) for support and providing additional time to be dedicated to research. YT thanks the National Institutes of Health for support (via grant number GM 069773). MM gratefully acknowledges the support of Nvidia Corporation with the donation of the Titan V GPU used for this research. L.M.F thanks the MINECO (CTQ2016-76061-P) project for financial support and the Centro de Computación Científica UAM (CCC-UAM) for computing time.

References

(1). Shin D; Sinkeldam RW; Tor Y Emissive RNA Alphabet. J. Am. Chem. Soc 2011, 133 (38), 14912–14915. 10.1021/ja206095a. [PubMed: 21866967]

(2). Park S; Otomo H; Zheng L; Sugiyama H Highly Emissive Deoxyguanosine Analogue Capable of Direct Visualization of B–Z Transition. Chem. Commun 2014, 50 (13), 1573 10.1039/c3cc48297a.

(3). Sholokh M; Sharma R; Shin D; Das R; Zaporozhets OA; Tor Y; Mély Y Conquering 2-Aminopurine's Deficiencies: Highly Emissive Isomorphic Guanosine Surrogate Faithfully

Monitors Guanosine Conformation and Dynamics in DNA. J. Am. Chem. Soc 2015, 137 (9), 3185–3188. 10.1021/ja513107r. [PubMed: 25714036]

(4). Sholokh M; Improta R; Mori M; Sharma R; Kenfack C; Shin D; Voltz K; Stote RH; Zaporozhets OA; Botta M; Tor Y; Mély Y Tautomers of a Fluorescent G Surrogate and Their Distinct Photophysics Provide Additional Information Channels. Angew. Chem. Int. Ed 2016, 55 (28), 7974–7978. 10.1002/anie.201601688.

(5). Martinez-Fernandez L; Gavvala K; Sharma R; Didier P; Richert L; Segarra Martì J; Mori M; Mely Y; Improta R Excited-State Dynamics of Thienoguanosine, an Isomorphic Highly Fluorescent Analogue of Guanosine. Chem. – Eur. J 2019, 25 (30), 7375–7386. 10.1002/chem.201900677. [PubMed: 30882930]

(6). Didier P; Kuchlyan J; Martinez-Fernandez L; Gosset P; Léonard J; Tor Y; Improta R; Mély Y Deciphering the PH-Dependence of Ground- and Excited-State Equilibria of Thienoguanine. Phys. Chem. Chem. Phys 2020, 22 (14), 7381–7391. 10.1039/C9CP06931C. [PubMed: 32211689]

(7). Sholokh M; Sharma R; Grytsyk N; Zaghzi L; Postupalenko VY; Dziuba D; Barthes NPF; Michel BY; Boudier C; Zaporozhets OA; Tor Y; Burger A; Mély Y Environmentally Sensitive Fluorescent Nucleoside Analogues for Surveying Dynamic Interconversions of Nucleic Acid Structures. Chem. - Eur. J 2018, 24 (52), 13850–13861. 10.1002/chem.201802297. [PubMed: 29989220]

(8). Han JH; Yamamoto S; Park S; Sugiyama H Development of a Vivid FRET System Based on a Highly Emissive DG-DC Analogue Pair. Chem. - Eur. J 2017, 23 (31), 7607–7613. 10.1002/chem.201701118. [PubMed: 28411372]

(9). Han JH; Park S; Hashiya F; Sugiyama H Approach to the Investigation of Nucleosome Structure by Using the Highly Emissive Nucleobase th DG–TC FRET Pair. Chem. – Eur. J 2018, 24 (64), 17091–17095. 10.1002/chem.201803382. [PubMed: 30207401]

(10). Kilin V; Gavvala K; Barthes NPF; Michel BY; Shin D; Boudier C; Mauffret O; Yashchuk V; Mousli M; Ruff M; Granger F; Eiler S; Bronner C; Tor Y; Burger A; Mély Y Dynamics of Methylated Cytosine Flipping by UHRF1. J. Am. Chem. Soc 2017, 139 (6), 2520–2528. 10.1021/jacs.7b00154. [PubMed: 28112929]

(11). Li Y; Fin A; McCoy L; Tor Y Polymerase-Mediated Site-Specific Incorporation of a Synthetic Fluorescent Isomorphic G Surrogate into RNA. Angew. Chem. Int. Ed 2017, 56 (5), 1303–1307. 10.1002/anie.201609327.

(12). Shin D; Lönn P; Dowdy SF; Tor Y Cellular Activity of SiRNA Oligonucleotides Containing Synthetic Isomorphic Nucleoside Surrogates. Chem. Commun 2015, 51 (9), 1662–1665. 10.1039/C4CC08809C.

(13). Raap J; Dreef CE; van der Marel GA; van Boom JH; Hilbers CW Synthesis and Proton-NMR Studies of Oligonucleotides Containing an Apurinic (AP) Site. J. Biomol. Struct. Dyn 1987, 5 (2), 219–247. 10.1080/07391102.1987.10506391. [PubMed: 2856028]

(14). Cuniasse Ph.; Fazakerley GV; Guschlbauer W; Kaplan BE; Sowers LC The Abasic Site as a Challenge to DNA Polymerase. J. Mol. Biol 1990, 213 (2), 303–314. 10.1016/S00222836(05)80192-5. [PubMed: 2342108]

(15). Cline SD; Jones WR; Stone MP; Osheroff N DNA Abasic Lesions in a Different Light: Solution Structure of an Endogenous Topoisomerase II Poison †. Biochemistry 1999, 38 (47), 15500–15507. 10.1021/bi991750s. [PubMed: 10569932]

(16). Lhomme J; Constant J-F; Demeuninck M Abasic DNA Structure, Reactivity, and Recognition. Biopolymers 1999, 52, 65–83. [PubMed: 10898853]

(17). Ayadi L; Coulombeau C; Lavery R Abasic Sites in Duplex DNA: Molecular Modeling of Sequence-Dependent Effects on Conformation. Biophys. J 1999, 77 (6), 3218–3226. 10.1016/S00063495(99)77152-3. [PubMed: 10585943]

(18). Melhuish WH Quantum efficiencies of fluorescence of organic substances: effect of solvent and concentration of the fluorescent solute. J. Phys. Chem 1961, 65 (2), 229–235. 10.1021/j100820a009.

(19). Godet J; Kenfack C; Przybilla F; Richert L; Duportail G; Mély Y Site-Selective Probing of CTAR Destabilization Highlights the Necessary Plasticity of the HIV-1 Nucleocapsid Protein to

Chaperone the First Strand Transfer. Nucleic Acids Res. 2013, 41 (9), 5036–5048. 10.1093/nar/gkt164. [PubMed: 23511968]

(20). Livesey AK; Brochon JC Analyzing the Distribution of Decay Constants in Pulse-Fluorimetry Using the Maximum Entropy Method. Biophys. J 1987, 52 (5), 693–706. 10.1016/S0006-3495(87)83264-2. [PubMed: 19431708]

(21). Brochon J-C [13] Maximum Entropy Method of Data Analysis in Time-Resolved Spectroscopy In Methods in Enzymology; Elsevier, 1994; Vol. 240, pp 262–311. 10.1016/S0076-6879(94)40052-0. [PubMed: 7823835]

(22). Macke T; Case DA Modeling Unusual Nucleic Acid Structures In Molecular Modeling of Nucleic Acids; LEONTES NB and SANTALUCIA J Jr.; American Chemical Society: Washington, DC, 1997; pp 379–393.

(23). Salomon-Ferrer R; Case DA; Walker RC An Overview of the Amber Biomolecular Simulation Package: Amber Biomolecular Simulation Package. Wiley Interdiscip. Rev. Comput. Mol. Sci 2013, 3 (2), 198–210. 10.1002/wcms.1121.

(24). Roe DR; Cheatham TE PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J. Chem. Theory Comput 2013, 9 (7), 3084–3095. 10.1021/ct400341p. [PubMed: 26583988]

(25). Miller BR; McGee TD; Swails JM; Homeyer N; Gohlke H; Roitberg AE *MMPBSA.Py*: An Efficient Program for End-State Free Energy Calculations. J. Chem. Theory Comput 2012, 8 (9), 3314–3321. 10.1021/ct300418h. [PubMed: 26605738]

(26). Dapprich S; Komáromi I; Byun KS; Morokuma K; Frisch MJ A New ONIOM Implementation in Gaussian98. Part I. The Calculation of Energies, Gradients, Vibrational Frequencies and Electric Field Derivatives. J. Mol. Struct. THEOCHEM 1999, 461–462, 1–21. 10.1016/S0166-1280(98)00475-8.

(27). Frisch MJ; Trucks W; Schlegel B Gaussian 09, Revision A.02 2009.

(28). Zhao Y; Schultz NE; Truhlar DG Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions. J. Chem. Theory Comput 2006, 2 (2), 364–382. 10.1021/ct0502763. [PubMed: 26626525]

(29). Zhao Y; Truhlar DG Density Functionals with Broad Applicability in Chemistry. Acc. Chem. Res 2008, 41 (2), 157–167. 10.1021/ar700111a. [PubMed: 18186612]

(30). Improta R; Santoro F; Blancafort L Quantum Mechanical Studies on the Photophysics and the Photochemistry of Nucleic Acids and Nucleobases. Chem. Rev 2016, 116 (6), 3540–3593. 10.1021/acs.chemrev.5b00444. [PubMed: 26928320]

(31). Cornell WD; Cieplak P; Bayly CI; Gould IR; Merz KM; Ferguson DM; Spellmeyer DC; Fox T; Caldwell JW; Kollman PA A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. J. Am. Chem. Soc 1995, 117 (19), 5179–5197. 10.1021/ja00124a002.

(32). Miertuš S; Scrocco E; Tomasi J Electrostatic Interaction of a Solute with a Continuum. A Direct Utilizaion of AB Initio Molecular Potentials for the Prevision of Solvent Effects. Chem. Phys 1981, 55 (1), 117–129. 10.1016/0301-0104(81)85090-2.

(33). Tomasi J; Mennucci B; Cammi R Quantum Mechanical Continuum Solvation Models. Chem. Rev 2005, 105 (8), 2999–3094. 10.1021/cr9904009. [PubMed: 16092826]

(34). Improta R; Barone V; Scalmani G; Frisch MJ A State-Specific Polarizable Continuum Model Time Dependent Density Functional Theory Method for Excited State Calculations in Solution. J. Chem. Phys 2006, 125 (5), 054103 10.1063/1.2222364. [PubMed: 16942199]

(35). Improta R; Scalmani G; Frisch MJ; Barone V Toward Effective and Reliable Fluorescence Energies in Solution by a New State Specific Polarizable Continuum Model Time Dependent Density Functional Theory Approach. J. Chem. Phys 2007, 127 (7), 074504 10.1063/1.2757168. [PubMed: 17718617]

(36). Martinez-Fernandez L; Improta R Photoactivated Proton Coupled Electron Transfer in DNA: Insights from Quantum Mechanical Calculations. Faraday Discuss. 2018, 207, 199–216. 10.1039/C7FD00195A. [PubMed: 29369316]

(37). Cuervo A; Dans PD; Carrascosa JL; Orozco M; Gomila G; Fumagalli L Direct Measurement of the Dielectric Polarization Properties of DNA. Proc. Natl. Acad. Sci 2014, 111 (35), E3624–E3630. 10.1073/pnas.1405702111. [PubMed: 25136104]

(38). Middleton CT; de La Harpe K; Su C; Law YK; Crespo-Hernández CE; Kohler B DNA Excited-State Dynamics: From Single Bases to the Double Helix. Annu. Rev. Phys. Chem 2009, 60 (1), 217–239. 10.1146/annurev.physchem.59.032607.093719. [PubMed: 19012538]

(39). Kadhane U; Holm AIS; Hoffmann SV; Nielsen SB Strong Coupling between Adenine Nucleobases in DNA Single Strands Revealed by Circular Dichroism Using Synchrotron Radiation. Phys. Rev. E 2008, 77 (2), 021901 10.1103/PhysRevE.77.021901.

(40). Cantor CR; Schimmel PR Techniques for the Study of Biological Structure and Function; Their Biophysical chemistry; pt. 2; W. H. Freeman: San Francisco, 1980.

(41). Lakowicz JR Principles of Fluorescence Spectroscopy, 3rd ed.; Springer: New York, 2006.

(42). Lomzov AA; Gorelov VV; Golyshev VM; Abramova TV; Pyshnyi DV 139 Analysis of Structure and Thermodynamics of Modified DNA Duplexes Using Molecular Dynamics Simulation. J. Biomol. Struct. Dyn 2015, 33 (sup1), 90–91. 10.1080/07391102.2015.1032772.

(43). Markham NR; Zuker M DINAMelt Web Server for Nucleic Acid Melting Prediction. Nucleic Acids Res. 2005, 33 (Web Server), W577–W581. 10.1093/nar/gki591. [PubMed: 15980540]

(44). Chen J; Dupradeau F-Y; Case DA; Turner CJ; Stubbe J Nuclear Magnetic Resonance Structural Studies and Molecular Modeling of Duplex DNA Containing Normal and 4'-Oxidized Abasic Sites †, ‡. Biochemistry 2007, 46 (11), 3096–3107. 10.1021/bi6024269. [PubMed: 17323932]

(45). Somsen OJG; van Hoek A; van Amerongen H Fluorescence Quenching of 2-Aminopurine in Dinucleotides. Chem. Phys. Lett 2005, 402 (1–3), 61–65. 10.1016/j.cplett.2004.11.122.

(46). Sinkeldam RW; Greco NJ; Tor Y Fluorescent Analogs of Biomolecular Building Blocks: Design, Properties, and Applications. Chem. Rev 2010, 110 (5), 2579–2619. 10.1021/cr900301e. [PubMed: 20205430]

(47). Jones AC; Neely RK 2-Aminopurine as a Fluorescent Probe of DNA Conformation and the DNA–Enzyme Interface. Q. Rev. Biophys 2015, 48 (2), 244–279. 10.1017/S0033583514000158. [PubMed: 25881643]

(48). Dziuba D; Postupalenko VY; Spadafora M; Klymchenko AS; Guérineau V; Mély Y; Benhida R; Burger A A Universal Nucleoside with Strong Two-Band Switchable Fluorescence and Sensitivity to the Environment for Investigating DNA Interactions. J. Am. Chem. Soc 2012, 134 (24), 10209–10213. 10.1021/ja3030388. [PubMed: 22591455]

(49). Seidel CAM; Schulz A; Sauer MHM Nucleobase-Specific Quenching of Fluorescent Dyes. 1. Nucleobase One-Electron Redox Potentials and Their Correlation with Static and Dynamic Quenching Efficiencies. J. Phys. Chem 1996, 100 (13), 5541–5553. 10.1021/jp951507c.

(50). Crespo-Hernández CE; Close DM; Gorb L; Leszczynski J Determination of Redox Potentials for the Watson−Crick Base Pairs, DNA Nucleosides, and Relevant Nucleoside Analogues. J. Phys. Chem. B 2007, 111 (19), 5386–5395. 10.1021/jp0684224. [PubMed: 17447808]

(51). Strickler SJ; Berg RA Relationship between Absorption Intensity and Fluorescence Lifetime of Molecules. J. Chem. Phys 1962, 37 (4), 814–822. 10.1063/1.1733166.

(52). Takaya T; Su C; de La Harpe K; Crespo-Hernandez CE; Kohler B UV Excitation of Single DNA and RNA Strands Produces High Yields of Exciplex States between Two Stacked Bases. Proc. Natl. Acad. Sci 2008, 105 (30), 10285–10290. 10.1073/pnas.0802079105. [PubMed: 18647840]

(53). Improta R; Barone V Interplay between "Neutral" and "Charge-Transfer" Excimers Rules the Excited State Decay in Adenine-Rich Polynucleotides. Angew. Chem. Int. Ed 2011, 50 (50), 12016–12019. 10.1002/anie.201104382.

(54). Wypijewska del Nogal A; Füchtbauer AF; Bood M; Nilsson JR; Wranne MS; Sarangamath S; Pfeiffer P; Rajan VS; El-Sagheer AH; Dahlén A; Brown T; Grøtli M; Wilhelmsson LM Getting DNA and RNA out of the Dark with 2CNqA: A Bright Adenine Analogue and Interbase FRET Donor. Nucleic Acids Res. 2020, gkaa525 10.1093/nar/gkaa525.

(55). Xu W; Chan KM; Kool ET Fluorescent Nucleobases as Tools for Studying DNA and RNA. Nat. Chem 2017, 9 (11), 1043–1055. 10.1038/nchem.2859. [PubMed: 29064490]

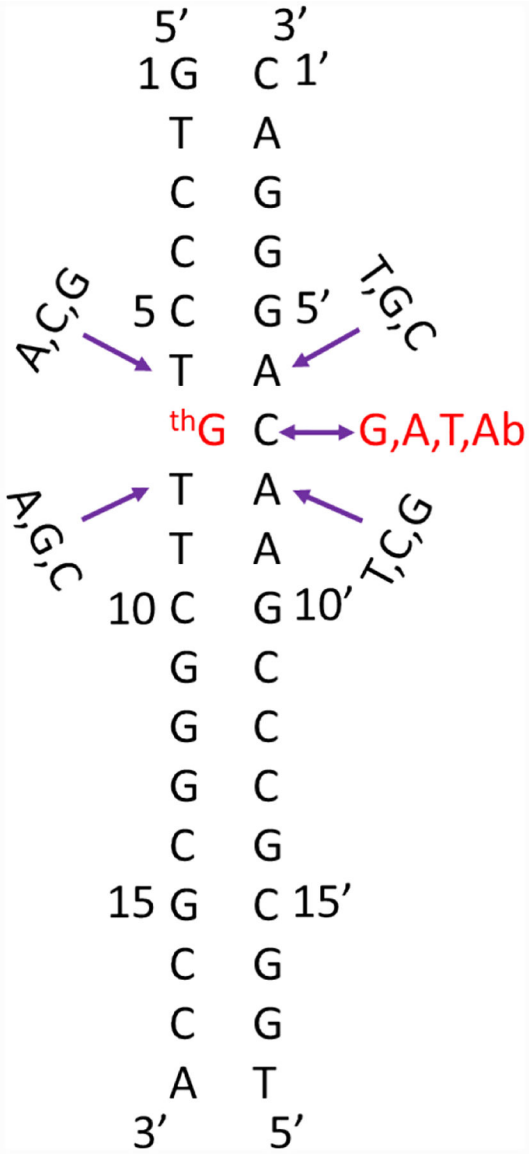


Figure 1:

Sequences of the thG-labeled (−)/(+)PBS duplexes used in this study. In all duplexes, the G7 residue of (−)PBS was replaced by the fluorescent analogue thG. To investigate the sequence context dependence of the fluorescence properties of thG, its flanking and opposite bases were systematically varied. Ab designates an abasic site, which corresponds to a site where the base is missing, but the sugar is present. The natural flanking T residues were systematically changed by two similar residues, and the opposite bases were concomitantly replaced by their matched Watson-Crick residues to minimize duplex destabilization. The duplexes were named using an XthGX/YZY triplet code comprising thG and its flanking residues and their opposite nucleobases. For example, the wild-type (−)/(+)PBS duplex is encoded as the TthGT/ACA duplex.

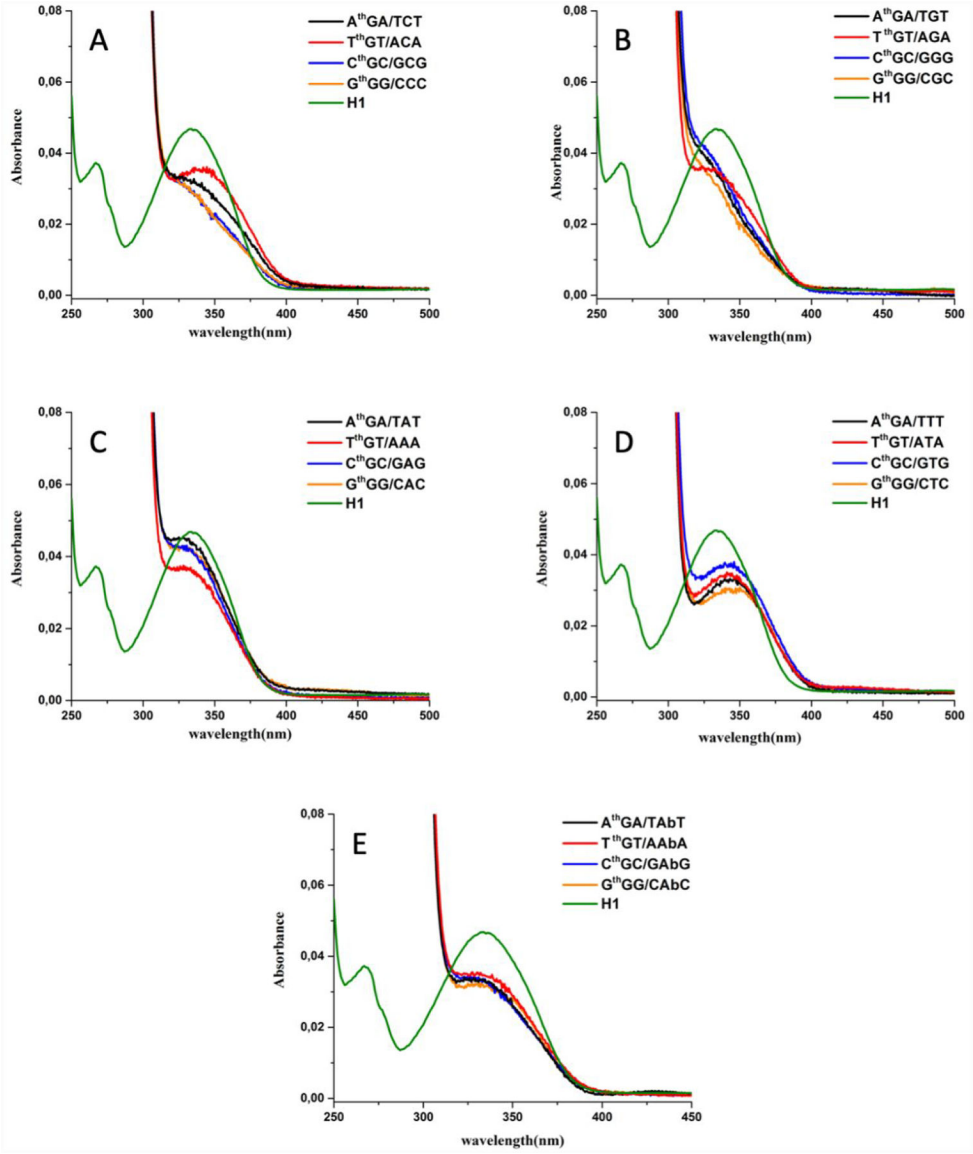


Figure 2:

Absorption spectra of thG-labelled matched (A) and mismatched (B-E) (+)/(−)PBS duplexes. In the mismatched duplexes, thG was opposite G (B), A (C), T (D) or an abasic site (E). In each panel, the neighboring residues of thG were systematically varied. The concentration of all duplexes was 10 μM. The absorption spectrum of the H1 tautomer of free thG at the same concentration in water is included for comparison (green spectra). All experiments were done in 25 mM TRIS-HCl buffer (pH=7.5), 30 mM NaCl and 0.2 mM $MgCl_2$.

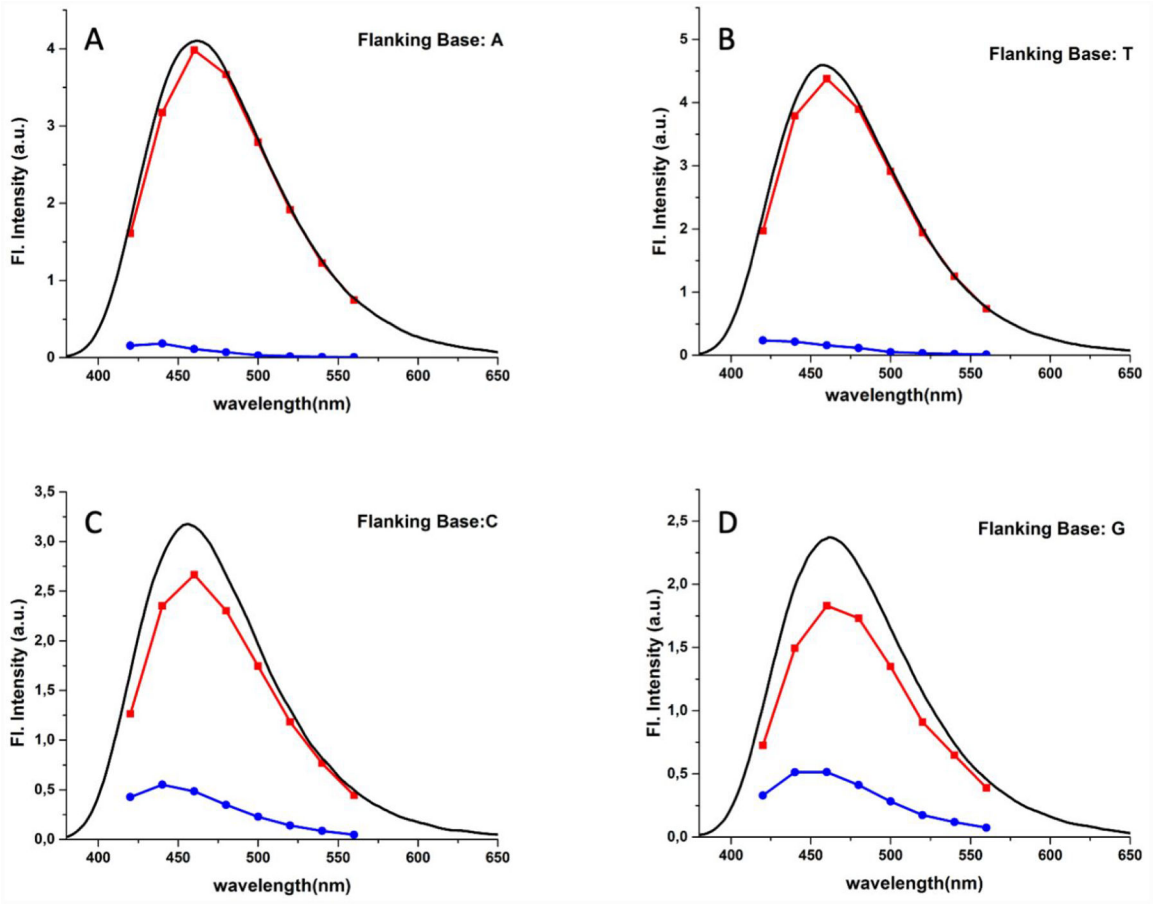


Figure 3:

Deconvolution of the fluorescence spectra of thG-labelled matched (+)/(−)PBS duplexes into two spectra associated with the long-lived lifetime component (red) and short-lived lifetime component (blue). The black lines represent the steady-state fluorescence spectra. Excitation wavelength was 360 nm.

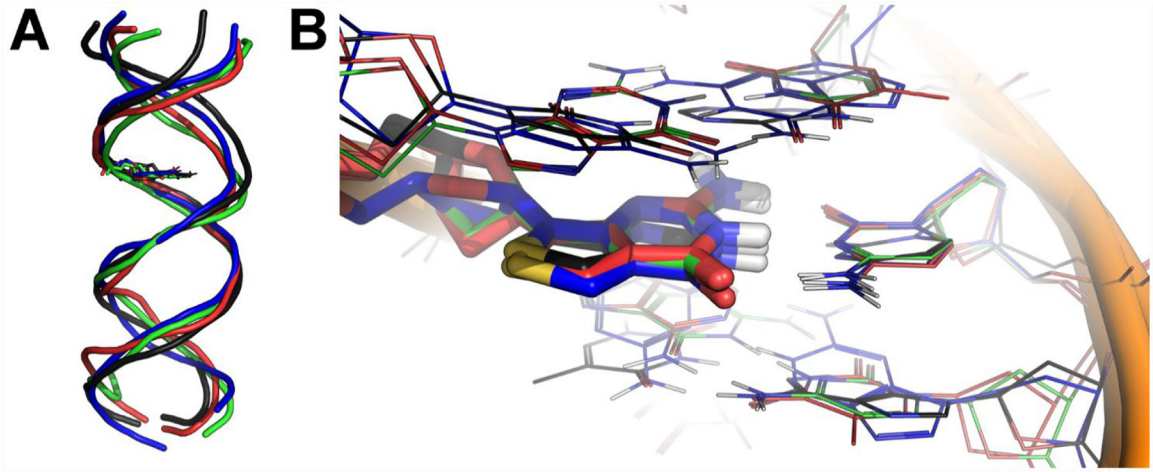


Figure 4:

Conformational stability of the matched thG-labelled (+)/(−)PBS duplexes in MD simulations. In the duplexes, thG was flanked either by T (black), A (red), C (blue) or G (green) residues. **A)** Structural superimposition of the most representative MD frames of (+)/(−)PBS matched duplexes. Only the cartoon representation of the phosphate backbone is shown, thG is shown as sticks. **B)** Magnification of the interaction of the thG-C base pair with its flanking base pairs. thG is shown as sticks, other nucleotides are shown as lines while the phosphate backbone is shown as a cartoon.

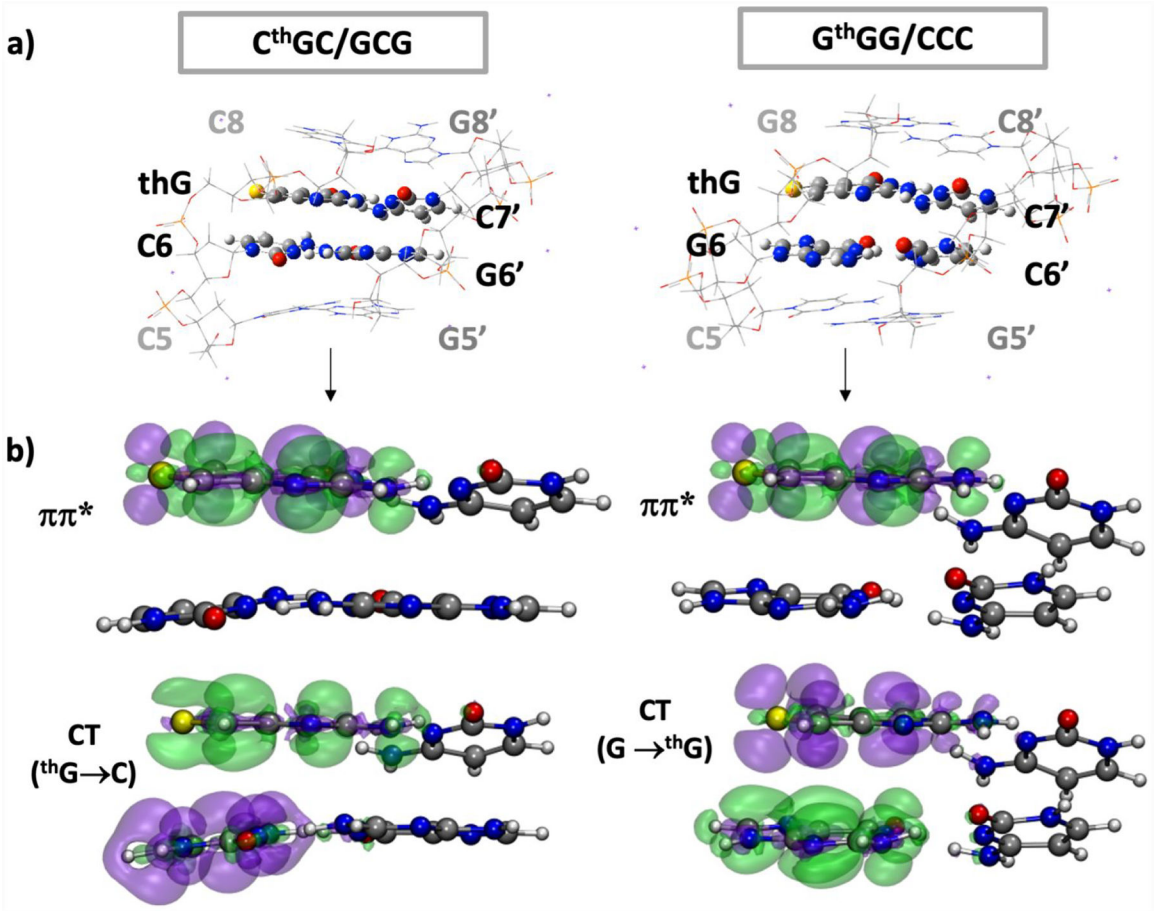


Figure 5: Ground-state geometries (A) and excited-state electronic densities (B) of the matched duplexes.

A) Most relevant geometries for the optimized minima at the QM(PCM/M052X/6–31G(d))/MM level of theory. Nucleobases treated at the QM level are in ball and sticks, whereas the MM system is depicted in lines. Numbering of residues is as in Figure 1. **B)** Excited-state density difference (S_n-S_0), green means negative difference (i.e from where the electron moves) while purple indicates more density (i.e where the electron moves to).

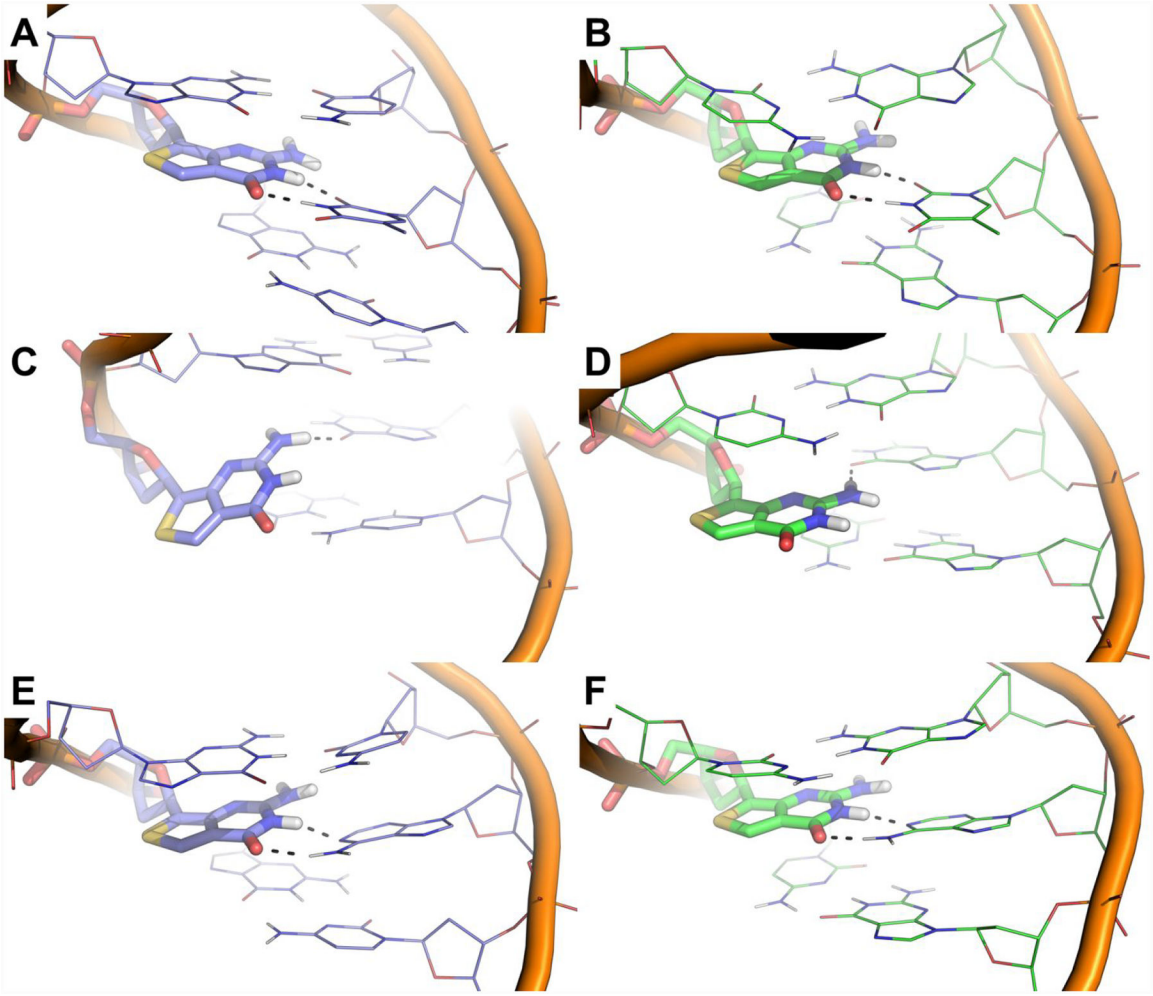


Figure 6. Conformational analysis of mismatched (+)/(−)PBS duplexes where thG is flanked by G and C bases.

A) Magnification of the interaction of the thG-T base pair with flanking G residues in the GthGG/CTC duplex. **B)** Magnification of the interaction of the thG-T base pair with flanking C residues in the CthGC/GTG duplex. **C)** Magnification of the interaction of the thG-G base pair with flanking G residues in the GthGG/CGC duplex. **D)** Magnification of the interaction of the thG-G base pair with flanking C residues in the CthGC/GGG duplex. **E)** Magnification of the interaction of the thG-A base pair with flanking G residues in the GthGG/CAC duplex. **F)** Magnification of the interaction of the thG-A base pair with flanking C residues in the CthGC/GAG duplex. thG is shown as sticks, other nucleotides are shown as lines while the phosphate backbone is shown as cartoon. H-bonds are shown as dashed lines. Mismatched duplexes where thG is flanked by G are coloured blue, while systems bearing thG flanked by C are coloured green.

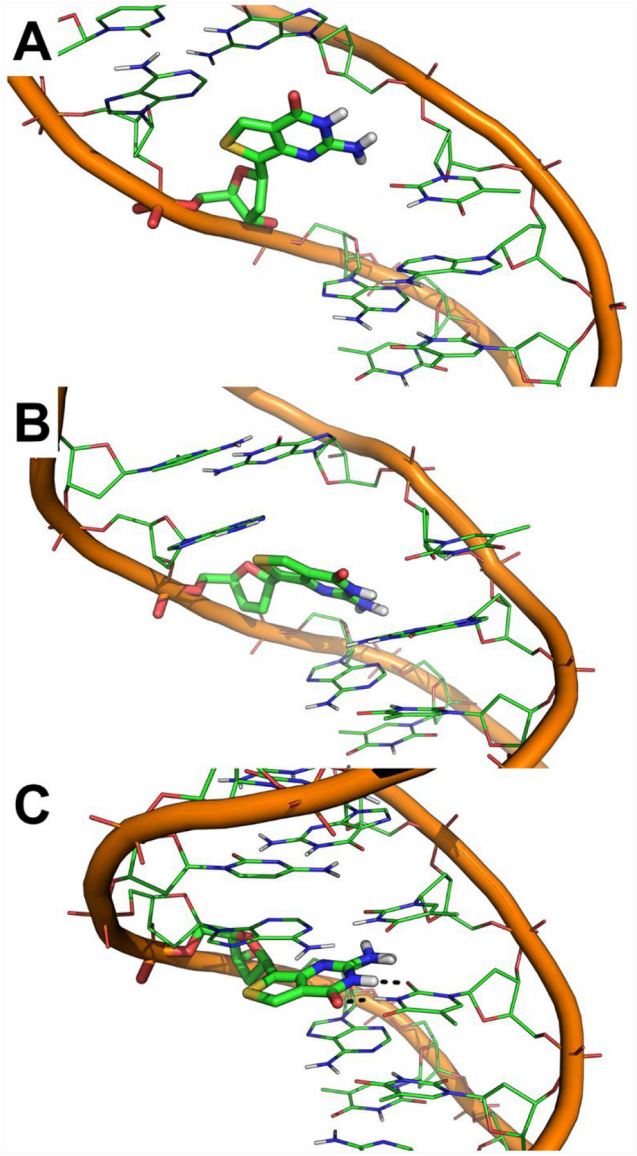


Figure 7.

Conformational behaviour of thG in the AthGA/TAT (A), AthGA/TGT (B) and AthGA/TTT (C) (−)/(+)PBS duplexes. Representative structures from MD simulations are shown. thG is shown as sticks, surrounding nucleotides as lines and the phosphate backbone is shown as cartoon. H-bonds are highlighted by black dashed lines. thG is shown as sticks, other nucleotides are shown as lines.

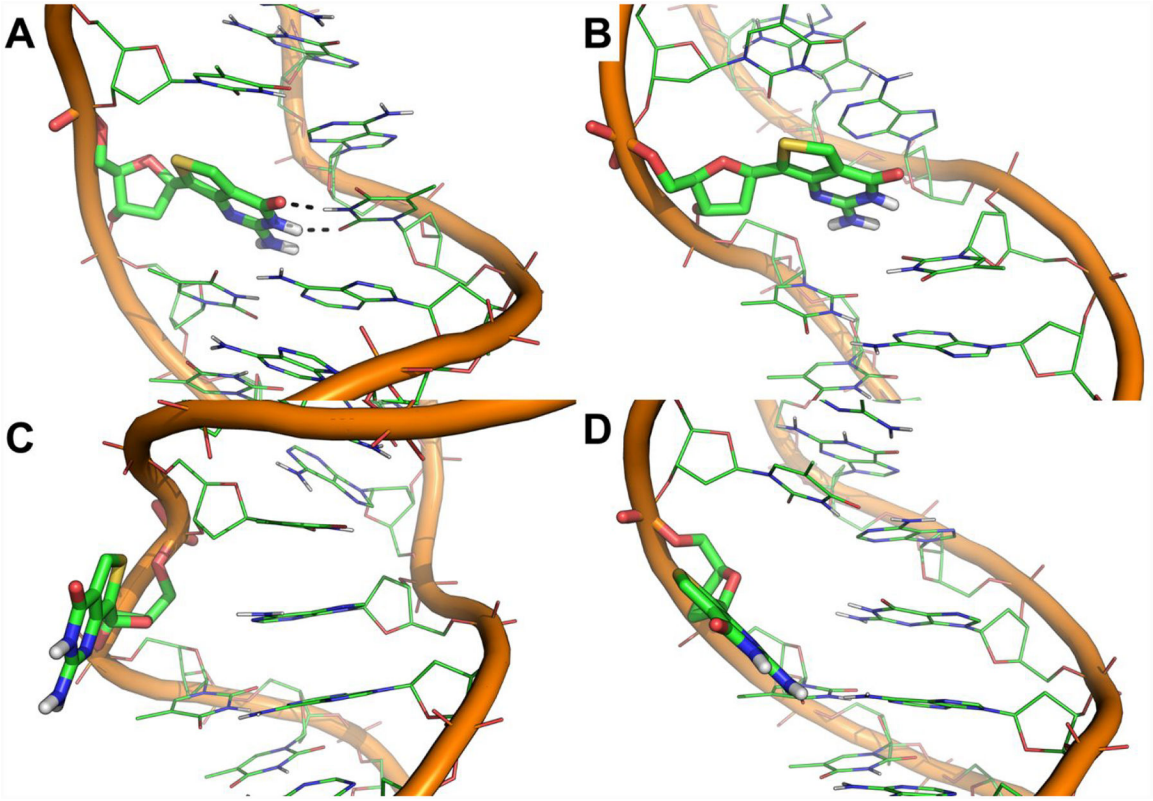


Figure 8.

Conformations of thG flanked with T residues, as obtained by MD simulations. **A)** The canonical Wobble-like thG-T base pair accounts for around 70% of MD frames (group 1). **B)** Destabilized conformation of the thG-T base pair, which accounts for around 30% of MD frames (group 2). **C)** Destabilized conformation of the thG-A mismatch; **D)** Destabilized conformation of the thG-G mismatch. thG is shown as sticks, nucleotides are shown as lines. H-bonds within the thG-T base pair are highlighted by black dashed lines.

Table 1:

Photophysical data of thG labelled (−)/(+) PBS matched duplexes

(−)PBS	(+)PBS	Hypochromism (%)[a]	QY[b]	τ_1 (ns)[c]	α_1[d]	f_1[e]	τ_2 (ns)[c]	α_2[d]	f_2[e]	<τ> (ns)	k_r (10^7×S^{-1})	k_{nr} (10^7×S^{-1})
AthGA	TCT	32	0.18	1.7	0.17	0.03	11.9	0.83	0.97	10.1	1.8	8.1
TthGT	ACA	29	0.15	2.4	0.14	0.04	11.1	0.86	0.96	9.9	1.5	8.6
CthGC	GCG	41	0.16	3.4	0.31	0.13	10.6	0.69	0.87	8.4	1.9	10.0
GthGG	CCC	41	0.15	4.4	0.39	0.19	12.3	0.61	0.81	9.2	1.6	9.2
thG H1 in water[f]			0.51				20.5				2.49	2.39
thG H1 in MeOH[f]			0.42				14.5				2.9	4.0

Standard Deviation (SD)= [a] ± 2%; [b] ± 0.02; [c] ± 0.1 — 0.3 ns; [d] ± 0.01 — 0.05. [e]Data from Martinez-Fernandez et al, 2019[5]. Excitation wavelength was 360 nm. The amplitudes, α_i, are calculated from the integrated areas under the DAS of each lifetime component normalized with respect to the total emitted intensity. The fractional intensities were calculated by $f_i = \alpha_i\tau_i/\langle\tau\rangle$. The radiative and non-radiative rate constants were calculated by: $k_r = QY/\langle\tau\rangle$ and $k_{nr} = 1/\langle\tau\rangle - k_r$, respectively.

Table 2.

Thermodynamic stability of thG-labeled (−)/(+)PBS duplexes.

	(−)PBS	(+)PBS	ΔEb (kcal/mol) ± SEM[a]	Tm[e]	
				Calc.[e]	Exp.[f]
Matched	GthGG	CCC	−75.0 ± 0.6	74.2	75.5
	CthGC	GCG	−71.2 ± 0.7	74.3	
	AthGA	TCT	−63.1 ± 0.6	68.2	70
	TthGT	ACA	−71.3 ± 0.8[b]	69	68.5
Mismatched	GthGG	CTC	−66.8 ± 0.7	68.6	
		CAC	−73.1 ± 0.8	69.3	68
		CGC	−61.5 ± 0.7	67.5	
		CAbC	−64.3 ± 0.8[c]	71.1	65.5
	CthGC	GTG	−65.6 ± 0.6	67.6	68
		GAG	−65.7 ± 0.8	64.9	66
		GGG	−59.4 ± 1.0	67.3	
		GAbG	−65.2 ± 1[c]	70.9	64
	AthGA	TTT	−56.2 ± 0.7	60.4	
		TAT	−48.7 ± 0.9	62.8	
		TGT	−51.1 ± 0.7	62.7	62
		TAbT	−55.8 ± 0.7[c]	68.7	60
	TthGT	ATA	−59.2 ± 1.0[d]	62.2	63
		AAA	−46.5 ± 0.7	62.4	
		AGA	−48.7 ± 1.0	63.1	
		AAbA	−40.2 ± 1.2[c]	68.9	

[a] Delta energy of binding, ΔEb, of thG-labeled (−)PBS to (+)PBS was calculated by the MM-PBSA approach along MD trajectories.

[b] Taken from our previous work[4].

[c] ΔEb was calculated by considering explicitly the 15 water molecules closest to thG in MD trajectories.

[d] The value refers to the most populated cluster (first group, around 70% of MD frames), while the second group (around 30% of MD frames) was estimated with a theoretical affinity of −51 ± 1 kcal/mol.

[e] The melting temperatures of the unlabeled duplexes were predicted by using the DINAMelt web server (http://unafold.rna.albany.edu/). They correspond to the inflection point of the predicted melting absorbance curves, using the experimental PBS (2 μM) and salt (30 mM NaCl and 0.2 mM $MgCl_2$) concentrations.

[f] Experimental melting temperatures in brackets were determined on the thG-labelled duplexes from the first derivative of the melting curves monitored from the absorbance changes at 260 nm (Fig. S6). The experimental *Tm* values were given for n = 2 as means +/−0.5 °C.

Table 3:

Absorption energies.

Vertical Absorption Energy (VAE) and oscillator strength *f* of the lowest energy excited-state (S_1), corresponding to a $\pi^{th}G) \rightarrow \pi^*(^{th}G)$ transition in the FC region. Values were obtained from LR-PCM/TD-M052X/6–31G(d) excited-state calculations on the structures extracted from QM/MM geometry optimizations. Hypochromism is estimated as the relative difference in % between the oscillator strength computed for free thG H1 tautomer and the one for the duplex.

(−)PBS	(+)PBS	VAE (eV)	VAE (nm)	*f*	*Hypochromism*
AthGA	TCT	4.30	288	0.11	27%
TthGT	ACA	4.31	287	0.11	27%
CthGC	GCG	4.24	292	0.11	27%
GthGG	CCC	4.37	283	0.11	27%
TthGT	ATA	^{g1}4.43	279	0.10	33%
		^{g2}4.59	270	0.12	20%
GthG**A**	CTC	4.24	292	0.07	53%
AthGA	TAT	4.63	268	0.16	0%
TthGT	AAA	^{g1}4.44	279	0.07	53%
		^{g2}4.56	272	0.10	33%
		^{g3}4.61	269	0.14	<1%
TthGT	CGC	4.56	272	0.16	0%
Free thG water[a]		4.55	273	0.15	
thG+2H_2O[b]		4.44	279	0.14	

[a] Bulk solvent effects are included by PCM.

[b] In addition to bulk solvent, two explicit water molecules were added.

Table 4.
Emission Energies of the matched and mismatched duplexes.

Vertical Emission Energies (VEE) and oscillator strengths (f) at the QM(PCM/M052X/6–31G(d) Linear Response (LR) level of theory are compared to those computed at the State Specific (SS)-PCM. Charge Transfer character (CT in a.u), radiative lifetime (τ_R in ns) and computed fluorescence lifetimes ($^c\langle\tau\rangle$ in ns) of the different minima for the different duplexes.

	(−)PBS	(+)PBS	Description.	^{LR}VEE (eV)	^{LR}f	^{LR}CT	^{SS}VEE (eV)	^{SS}f	$^{SS}\tau_R$	$^c\langle\tau\rangle$
Match	AthGA	TCT	Min-ππ*	3.28	0.07	0.03	3.20	0.05	74	13.4
			Min-CT (thG→A)	3.77	0.01	0.58	2.53	0.00		
	TthGT	ACA	Min-ππ*$_{model1}$	3.43	0.08	0.01	3.32	0.05	64	9.7
			Min-CT$_{model2}$ (thG→T)	3.34	0.00	0.86	1.68	0.00		
	CthGC	GCG	Min-ππ*	3.42	0.08	0.01	3.33	0.06	52	8.3
			Min-CT (thG→C)	[a]3.88	0.03	0.72	2.10	0.01		
	GthGG	CCC	Min-ππ*	3.47	0.08	0.03	3.33	0.05	64	9.6
			Min-CT (G→thG)	3.15	0.02	−0.22	2.80	0.01		
Mismatched	GthGG	CTC	Min-ππ*	b						
			Min-CT (G→thG)	4.24	0.07	−0.22	2.91	0.01	7000	
	AthGA	TAT	Min-ππ*	3.67	0.11	0.03	3.39	0.08	29	14.0
			Min-CT	c						
	TthGT	ATA	^{g1}Min-ππ*	3.46	0.07	0.04	3.32	0.05	68	23.7
			^{g1}Min-ct (thG→T)	3.36	0.06	0.06	3.24	0.04		
	TthGT	AAA	^{g3}Min-ππ*	3.50	0.08	0.01	3.29	0.06	62	22.8
			^{g1}Min-ct (thG→T)	2.93	0.01	0.16	2.80	0.01		
	TthGT	AGA	Min-ππ*	3.63	0.11	0.00	3.33	0.08	28	2.8
			Min-CT	c						
thG+water			Min-ππ*	3.60	0.11	-	3.22	0.08	31	15.8
thG+2H$_2$O			Min-ππ*	3.45	0.11	-	3.15	0.08	31	15. 8

Notes: a) No real minimum but a low energy gradient region was found, b) no clear ππ * minimum was found, c) no clear CT minimum was located. The radiative lifetime (τ_R) was estimated by using a simplified version of the Strickler-Berg formula: $^{Theor}\tau_R = \frac{1}{k_R}$ with $k_R = \frac{4\Delta E^3}{3\,c^3}\mu^2$, where ΔE is the vertical emission energy and the dipole moment of the transition. The computed lifetime is calculated by: $^c\langle\tau\rangle = k_R \times QY$, where QY is the experimental quantum yield. Further details on the calculations are given in section 4 of SI.

Table 5:

Photophysical properties of thG-labelled (−)/(+) PBS mismatched duplexes

(−)PBS	(+)PBS	Hypochr omism (%)	QY	τ_1 (ns)	α_1	f_1	τ_2 (ns)	α_2	f_2	$\langle\tau\rangle$ (ns)	k_r ($10^7\times S^{-1}$)	k_{nr} ($10^7\times S^{-1}$)
GthGG	CTC	37	0.10	2.4	0.43	0.16	9.3	0.57	0.84	6.3	1.6	14.2
GthGG	CAC	13	0.17	4.4	0.39	0.18	12.9	0.61	0.82	9.6	1.8	8.7
GthGG	CGC	32	0.23	2.5	0.29	0.05	19.8	0.71	0.95	14.8	1.5	5.2
GthGG	CAbC	28	0.15	4.0	0.49	0.25	11.6	0.51	0.75	7.9	1.9	10.8
CthGC	GTG	24	0.15	2.1	0.42	0.10	13.5	0.58	0.90	8.7	1.7	9.8
CthGC	GAG	15	0.34	2.7	0.19	0.03	18.8	0.81	0.97	15.7	2.2	4.2
CthGC	GGG	30	0.21	3.5	0.26	0.08	14.8	0.74	0.92	11.8	1.8	6.7
CthGC	GAbG	26	0.12	4.7	0.68	0.42	14.5	0.32	0.58	7.8	1.5	11.2
AthGA	TTT	33	0.37	2.3	0.31	0.04	27.3	0.69	0.96	19.6	1.9	3.2
AthGA	TAT	12	0.48	2.8	0.22	0.03	28.6	0.78	0.97	22.9	2.1	2.3
AthGA	TGT	28	0.42	2.3	0.17	0.02	26.9	0.83	0.98	22.7	1.8	2.6
AthGA	TAbT	28	0.47	2.4	0.18	0.02	25.0	0.82	0.98	20.9	2.2	2.5
TthGT	ATA	25	0.35	3.0	0.27	0.04	26.8	0.73	0.96	20.4	1.7	3.1
TthGT	AAA	25	0.42	3.0	0.24	0.04	23.8	0.76	0.96	18.8	2.2	3.1
TthGT	AGA	31	0.33	5.5	0.25	0.09	19.4	0.75	0.91	15.9	2.1	4.2
TthGT	AAbA	23	0.25	3.4	0.28	0.07	16.9	0.72	0.93	13.1	1.9	5.7

All reported values are the means for two to four experiments. The standard errors of the mean of the reported values are 8% for the QY, 10% for hypochromism, ± 0.1 — 0.3 ns for τ1, ± 0.2 — 0.8 ns for τ2, < 0.05 for the amplitudes (α_i) and fractional intensities (f_i). The radiative and non-radiative rate constants were calculated as described in Table 1.